DOCUMENT RESUME

ED 091 941 FL 006 110

AUTHOR	Bell, Alan
TITLE	Some Patterns of Occurrence and Formation of Syllable Structures. Working Papers on Language Universals, No. 6.
INSTITUTION	Stanford Univ., Calif. Committee on Linguistics.
PUB DATE	Oct 71
NOTE	115p.
EDRS PRICE	MF-$0.75 HC-$5.40 PLUS POSTAGE
DESCRIPTORS	Descriptive Linguistics; Language Classification; Language Typology; *Language Universals; Linguistic Patterns; *Linguistic Theory; Structural Analysis; *Syllables
IDENTIFIERS	*Markow Chain Process

ABSTRACT

A particular aspect of syllable structure, length of syllable margins, was investigated with the aid of a Markov chain model. The model represented explicitly the dynamic relationship between types of syllable structure and the historical processes that affect them. It is proposed that the regularities concerning syllable types (universality of CV syllables, etc.) are best understood and explained if they are assumed to be nondeterministic--referring to the relative frequency of occurrence rather than occurrence vs. nonoccurrence. A number of specific generalizations about other aspects of syllable structure are also proposed on the basis of the data and the analysis. (Author)

ED 091941

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

WORKING PAPERS ON LANGUAGE UNIVERSALS

Number Six

October 1971

Language Universals Project
Committee on Linguistics
Stanford University
Stanford, California

FL 006 110

PAGES 1-22 WERE [redacted] MISSING FROM THE DOCUMENT THAT WAS SUBMITTED TO ERIC DOCUMENT REPRODUCTION SERVICE.

FL 006110

SOME PATTERNS OF OCCURRENCE AND FORMATION OF SYLLABLE STRUCTURES

Alan Bell

ABSTRACT

A particular aspect of syllable structure, length of syllable margins, was investigated with the aid of a Markov chain model. The model represented explicitly the dynamic relationship between types of syllable structure and the historical processes that affect them. It is proposed that the regularities concerning syllable types (universality of CV syllables, etc.) are best understood and explained if they are assumed to be nondeterministic -- referring to the relative frequency of occurrence rather than occurrence vs. nonoccurrence. A number of specific generalizations about other aspects of syllable structure are also proposed on the basis of the data and the analysis.

Introduction. In this investigation, I sought generalizations about syllable structures, taking into account both their occurrence and formation. The universe of possible languages, which is in principle the scope of the generalizations, is conceived as composed of types or states related to each other by processes of change. This obliges one to consider properties of language simultaneously with the ways that they arise and are modified. Further discussion and illustration of the approach may be found in Greenberg (1960a, 1969).

1. Typologies of syllable structure

Syllables have an internal constituent structure, dominated by the opposition of margin to nucleus. Based on this opposition, four types of syllables can be defined, symbolized conventionally V, CV, VC, CVC. Depending on the number and nature of the segments in the margins and nuclei further subtypes may be defined, such as CCV, CVV, CVN, etc. It is the number of segments in the margins that I will mainly be concerned with. Some preliminary observations are offered in Section 2.4.3., and the reader can refer to Greenberg (1965) for generalizations about the kinds of marginal sequences. Indeed, my scope is even more limited. I practically ignore the distribution of longer clusters, finding much more intriguing patterns based on the occurrence of syllables with and without margins and with and without clusters (of whatever size).

1.1 Generalizations of Jakobson and Greenberg. Generalizations about the syllable structures of languages have focused on the privileged status of CV syllables. Jakobson and Halle (1956) state that

> since many languages lack syllables without a prevocalic consonant and/or with a postvocalic consonant, CV . . . is the only universal model of the syllable. (p. 37) . . . There are languages where every syllable consists of a consonant and a succeeding vowel (CV). . . (p. 20)

In other words, occurring syllable structures are those which have just the syllable type CV or else include the syllable type CV among others.

Greenberg later refined this generalization, modifying it slightly:

> ... every language has syllables of the form CV... in addition to whatever other type it may possess. ...if syllables containing sequences of n consonants in a language are to be found as syllabic types, then sequences of n-1 consonants are also to be found in the corresponding position (prevocalic or postvocalic) except that CV→ V does not hold. (Greenberg, Osgood and Jenkins 1966: xxv)

This implies that syllable structures containing just one syllable type are not to be found. A further implication is that syllable type VC does not occur unless the syllable type V also occurs and syllables with complex margins do not occur without those of lesser complexity. Syllable structures with simple margins are therefore limited to CV, V; CV, CVC; CV, V, VC; CV, V, CVC; and CV, V, VC, CVC -- 5 out of the 15 possible subsets of the set of four basic syllable types. (A set of n elements has 2^n subsets if the null set is counted. Here $2^4-1 = 15$.)

1.2 The subset typology of syllable structure. In the generalizations of Jakobson and Greenberg, languages are categorized according to their inventory of syllable types. This might be called the 'subset' typology of syllable structure, since every syllable structure is a subset of a basic set of syllable types. This simple system of categorization led to a clear statement of certain regularities of syllable structure, and has also been used by Martinet (1964) to investigate diachronic tendencies toward a structure of open syllables. It has not, however, led to any further understanding of the nature of the observed regularities. A new framework is needed for the present investigation, which seeks to extend what we know about the regularities of syllable structure and to explain them in terms of diachronic process.

1.3 Inadequacies of the subset typology. The subset typology is inadequate for our purposes, first of all, because it is too gross.

Languages differ greatly in the constraints that they impose on the combination of syllabic types, and there is reason to suppose that these differences may be significant. For example, ESKIMO and LUGANDA are both languages whose subset syllable structure is CV, V, VC, CVC. But in ESKIMO, the syllable types can appear virtually anywhere within a word, whereas there are considerable constraints in LUGANDA: V, VC syllables only occur initially, and VC, CVC may not occur finally. Secondly, and more seriously, it is inappropriate from the point of view of process. Formally, CV > CVC is a process, but to take such changes as unit processes is unrevealing. The phonetic processes that shape syllable structure are largely consonant and vowel loss, and to a lesser extent the epenthesis of consonants and vowels. I prefer to relate syllable structures not to processes such as CV > CVC, but rather to such processes as $V \rightarrow \emptyset$ / CVC__#, or V → -syllabic / CV__C, both of which usually have the effect of changing a former CV syllable into a CVC syllable.

1.4 A parametric typology of syllable structure. Both of these weaknesses can be traced to the subset typology's representation of the syllable as a disembodied unit. This convenient simplification, which treats syllables as abstracted from segment strings, fails to capture the role of the syllable as a unit of organization of strings of segments. The view of syllable structure as consisting of the ways that segment strings can be organized into syllables is a prerequisite to our deeper understanding of it.

We can construct a systematic characterization of the ways of organizing strings into syllables in the following manner. First, the domain over which syllable structure is defined is taken to be the next higher prosodic domain above the syllable, which in most cases corresponds to the phonological word. Next, the boundary conditions at the beginning and the end of words are specified in terms of the minimum

and maximum number of nonsyllabic segments occurring in each position. Then, treating all internal syllable junctions together, the minimum and maximum number of marginal segments between syllable nuclei are specified.[1]

As a first approximation, then, the syllable structure of a language is characterized by six parameters, the minimum and maximum number of nonsyllabic segments occurring in initial, medial, and final position. Their values range from 0 to n, where n is some bound on the longest possible string of nonsyllabic segments. This characterization and its variants will be referred to as the parametric typology of syllable structure. In the above form, it contains far too many categories for our purposes, numbering in the thousands, depending on the value of n. Some of the categories must be collapsed. Since my primary interest is not the complexity of syllable margins, syllable structures with complex margins are lumped together. Specifically, structures with initial and

---

[1] The conventions of Section A.3.2. are used in deciding upon the number of segments in a marginal sequence. The marginal segments are referred to henceforth in this chapter as consonants, and are symbolized by the letter C.

A rigorous and uniformly applicable definition of prosodic domains is notoriously difficult. While it is common for the domain of syllabic organization to coincide with an accentual domain, so that there is a well-defined phonological word, in many languages (FRENCH and LUGANDA, for example) the domain of syllabic organization extends over some larger unit such as the phrase. (This distinction has been carefully characterized by Pulgram (1969), who terms the two domains respectively nexus and cursus.) The situation is not accurately known for still other languages. I have dealt with this problem by insisting on a clearly medial environment for the determination of the minimum and maximum number of internal marginal segments, but to include phrase-terminal position for the determination of the initial and final minima and maxima. This means, for example, that a language with hiatus only at word boundary is said to lack hiatus (internal minimum of one marginal segment), and a language whose word-initial vowels disappear under sandhi except phrase-initially is counted as having word-initial vowels (initial minimum of one marginal segment).

final margins of two or more segments are placed together, as are those with medial marginal strings of three or more segments. Syllable structures are now represented by the parameters and values in Figure 1. The starred values of the parameters do not occur.

Parameter	Symbol	Values
Maximum initial margin sequence	$\bar{I}$	0, 1, 2 (= ≥2)
Minimum initial margin sequence	$\underline{I}$	0, 1, 2^*
Maximum medial margin sequence	$\bar{M}$	0, 1, 2, 3 (= ≥3)
Minimum medial margin sequence	$\underline{M}$	0, 1, 2^*, 3^*
Maximum final margin sequence	$\bar{F}$	0, 1, 2 (= ≥2)
Minimum final margin sequence	$\underline{F}$	0, 1, 2^*

Figure 1. Parametric typology of syllable structure.

The parameter values are obviously constrained by the condition that the maximum number of segments is no less than the minimum number of segments in any position, so that $\bar{I} \geq \underline{I}$, etc.

Certain values of the parameters, starred in Figure 1, are assumed not to occur. The categories with values $\underline{I} = 2$ of $\underline{F} = 2$ would contain languages with initial or final consonant clusters but no initial or final single consonants, in violation of Greenberg's (1965) generalizations on consonant clusters, which are also supported by the present data. Moreover, no language in the sample possesses medial clusters without also possessing simple intervocalic consonants. I conclude that $\underline{I}$, $\underline{M}$, $\underline{F}$ may take only the values 0 and 1.

1.5 Other aspects of syllable structure. The parametric typology provides a skeletal characterization of syllable structure. Other important aspects of syllable structure it does not account for are syllable division, word length, nucleus structure, margin subtypes, relation to accentual systems and positional constraints. These are briefly discussed

immediately below and in Section 2.4.

Although the parametric typology provides far more total information about syllable structure than the subset typology, the inventory of syllable types of a language is not always fully determined by its parameter values. The division or adherence of medial consonants and word length can also affect what syllable types a language has. Unless otherwise stated, the syllable types of a language in a given category of the parametric typology are determined under the assumption that VCV and VCCV sequences are uniformly divided as V.CV and VC.CV. The assumption does not always hold, of course. It is a convenient procedural fiction. For the languages of the sample, and perhaps in general, the syllable types are correctly determined even when the assumption does not hold. See Section 2.4.1. below.

I will also usually assume that a language has at least monosyllabic, disyllabic, and trisyllabic words. Stronger constraints on minimum or maximum word length sometimes exclude certain syllable types. The parameters $\overline{M}$ and $\underline{M}$ are undefined for monosyllabic languages, of course, and these languages will not be treated within in the typology.[2] Examples of the effect of word length are discussed in Section 2.4.2.

Another implicit assumption is that constraints on medial marginal sequences are uniform over the word. This is not always true. For example, in one dialect of the WESTERN DESERT language of Australia, medial -CC- sequences occur only after the first and second nuclei of a word (Douglas 1964). Such constraints can affect the syllable type inventory. Consider a language whose words begin with one consonant or none ($\overline{I} = 1$, $\underline{I} = 0$), end in one consonant or none ($\overline{F} = 1$, $\underline{F} = 0$), and have medial marginal sequences of one or two consonants ($\overline{M} = 2$, $\underline{M} = 1$).

---

[2] Whether there in fact do exist monosyllabic languages is not known. Trubetzkoy's (1969) classification of BURMESE as such is not in accord with the sandhi and stress phenomena that occur over multisyllabic sequences. CANTONESE, from the description of Cheng (1968), may be monosyllabic, and was accordingly not included in the basic sample.

It will usually be supposed to have syllables of types CV, V, VC, CVC. But say that no -CC- sequences occur between the first and second syllable. Then the syllable type VC cannot occur, since only initial syllables may have a nuclear onset.

Another implicit assumption is that the initial, medial, and final parameters are independent. That this is not strictly true, and is sometimes of consequence, may be seen from NORTH PALESTINIAN ARABIC. There a word may begin with vowel only if it is followed by two consonants in phrase-initial position (Blanc 1953). Since hiatus does not occur ($\underline{M} = 1$), this means that the syllable type V cannot occur even though VC does occur, a fact not directly predictable from the language's parametric category. The example is of interest since it is a counter-example to Greenberg's proposed generalization that the syllable type VC occurs only if the type V also occurs.

No attempt has been made to account for the structure of syllable nuclei. All nuclei with vowels are lumped together, whatever their internal structure.[3]

Syllables with consonantal nuclei are considered to be separate, peripheral types. This treatment is possible because such syllables are a marginal system, infrequent and readily treated apart, in almost all languages (with exceptions such as KORYAK). I also preferred to do this rather than treat all nuclei together, because syllabic consonants appear to behave differently from vocalic nuclei in certain important respects. For example, HOTTENTOT roots in CV, V have variants in CV?V, but roots in CV, N do not have CV?N variants (Beach 1938). Occasionally, some syllable types may occur only with consonantal nuclei, as in NAVAHO, whose only VC syllables have syllabic n for a peak. A

---

[3] Note that the arbitrary decision on the classification of possible satellite resonants (Section A. 3. 2.) means that possible cases of VR nuclei are treated as nucleus + margin.

systematic study of such distinctions would certainly help to discove more about the nature of syllabic consonants, but for the purposes of a study of the general nature of syllabic structure their inclusion would have introduced an unjustified complication.

2. Survey of types of syllable structure

The following survey of the types of syllable structure is based on a sample of 144 languages. The Appendix lists the languages, describes how they were chosen, and states some of the assumptions made in classifying them.

2.1 Some preliminary observations. No languages were found corresponding to the following parameter values: 1) $\bar{I} = 0$, 2) $\bar{M} = 0$, 3) $\underline{M} = 1$, and $\underline{F} = 1$. The lacunae may be expressed by the generalizations.

(1) All languages have initial margins.

(2) All languages have medial margins (That is, hiatus is not obligatory.)

(3) All languages have medial margins of one segment.

(4) All languages have some open final syllables.

A few languages are almost exceptions to (1), (3) and (4). The words of OYKANGAND, an Australian language, have been characterized as obligatorily beginning with a vowel (Sommer 1969). However a dozen or so common words have no initial vowel phrase-initially. Most words of OYKANGAND and related dialects also end in consonants, but in terms of our framework, there are many final open syllables, since, by the assumptions of Section A.3.2, diphthongs are not considered to constitute a phonetic margin. (In fact, there are final non-diphthongal nuclei, since /iy/ is phonetically [i·].) BUSHONG, a Bantu language of the Congo, has undergone extensive loss of final vowels, leading to a great predominance of words with closed syllables, but there

still remain numerous words with final vowels.

One might feel that the automatic gemination found in MADURESE (Stevens 1968) and in WIYOT (Teeter 1964) indicates that there is some posibility of complete gemination of all medial consonants in a language, violating (3). I judge the likelihood to be very remote. Gemination in MADURESE is limited to consonants after schwa, and in WIYOT to those after short vowels. The process thus appears to reflect a sort of harmony of syllabic heaviness, rather than any tendency toward total gemination. Glottal stop in MADURESE and the lateral flap in WIYOT are excepted from gemination.

2.2 General classification. A summary of the number of languages with each type of syllabic structure is presented in Figures 2, 3, and 4. The total for Figure 2 is greater than 144. In order to obtain the fullest possible picture of the range of occurrence of types, languages have been assigned here to several types where there is variation between syllable structures. The variation may be idiolectical, stylistic, or free. For example, IJO, a Niger-Congo language, has initial clusters as variants in words like [pɪrɪ] ~ [prɪ], and is classed as $\bar{I} = 1$ and $\bar{I} = 2$. In some cases the indigenous structure of a language and a newer structure due to loans are counted separately. TAGALOG, in which hiatus occurs only in loans, is an example.

In the distributions of Figures 3 and 4 each language or dialect is counted once. Where there was a choice of types, the innovating type was retained in the case of phonetic alternation; the type based on indigenous forms was retained if one classification was based on loans.

The languages in each type are listed in Figures 5a through 5f. Languages in parenthesis are alternant classifications not counted in Figures 3 and 4.

Values of the parameters for which there are no occurrences are not shown. Since for every language, $\underline{F} = 0$ (see generalization (3) above), this parameter was omitted entirely.

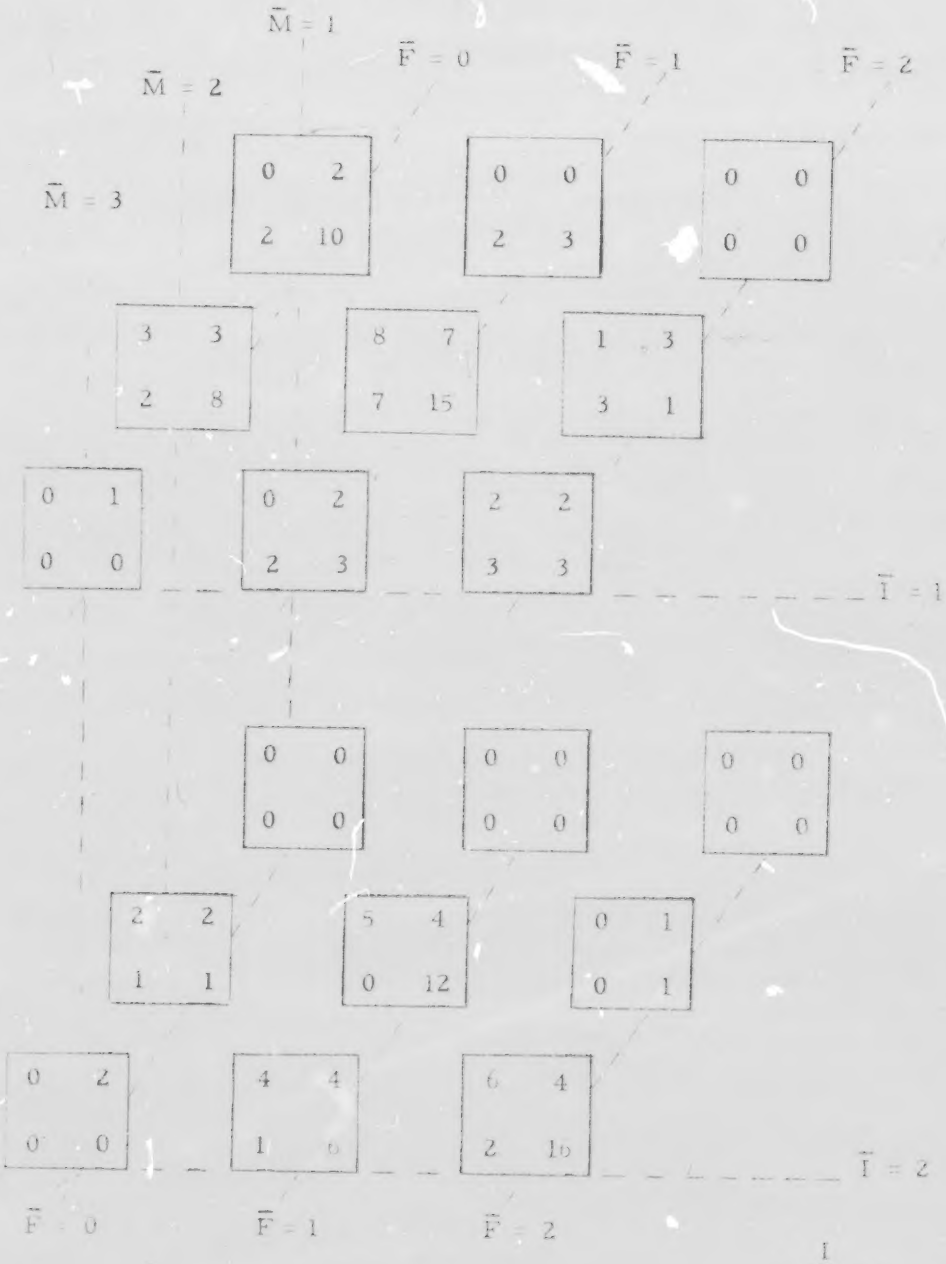
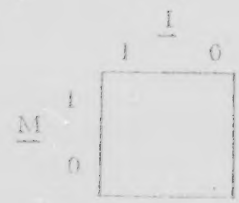


Figure 2. Number of languages in each syllable structure category. The quadrants of each square correspond to the values of $\underline{M}$ and $\underline{I}$ as in the key at right.



$\bar{M} = 1$	11	5	0	
$\bar{M} = 2$	14	32	6	$\bar{I} = 1$
$\bar{M} = 3$	1	2	8	

$\bar{M} = 1$	0	0	0	
$\bar{M} = 2$	6	20	2	$\bar{I} = 2$
$\bar{M} = 3$	1	14	22	

$\bar{F} = 0$ $\bar{F} = 1$ $\bar{F} = 2$

Figure 3. Number of languages in each cluster category.

	$\underline{I} = 1$	$\underline{I} = 0$
$\underline{M} = 1$	28	32
$\underline{M} = 2$	18	66

Figure 4. Number of languages in each minimum margin category.

	$\bar{M} = 1$	$\bar{M} = 2$	$\bar{M} = 3$
$\underline{I} = 1$, $\underline{M} = 1$		Chacobo	
		Colorado	
		Siona	
$\underline{I} = 1$, $\underline{M} = 0$	Ellice	Piro	
	(Tswana)	Southern Paiute	
$\underline{I} = 0$, $\underline{M} = 1$	Paresi	Ignaciano	Kannada
	(Swahili)	Luganda	
		Swahili (Bravanese)	
$\underline{I} = 0$, $\underline{M} = 0$	Auca	Campa	
	Balese	Dogon (Toro)	
	Esako	(Galla)	
	Fijian	(Ijo)	
	Kikuyu	Maori	
	(Maori)	Nukuoro	
	Swahili	Eastern Popoluca	
	Tswana	Secoya	
	Warao		
	Yoruba		

Figure 5a. Classification of languages by syllable structure.
$\bar{I} = 1$, $\bar{F} = 0$

	$\bar{M} = 1$	$\bar{M} = 2$	$\bar{M} = 3$
$\underline{I} = 1, \underline{M} = 1$		Bilin	Hakka
		Cofan	Khasi
		Fulani	Mandarin
		Gugu-Yalanji	Tibetan
$\underline{I} = 1, \underline{M} = 0$	Arabela	Dogon	Ifugao (Koya)
	Hottentot	(Kamba)	Korean (Hausa)
		Guarani	Navaho
		Hausa	Sangir
$\underline{I} = 0, \underline{M} = 1$		Berti	(Mongol) Mongol
		Burmese	Sentani Wik-Munkan
		Cayapa	Tagalog
		Western Desert (Oolde:)	
$\underline{I} = 0, \underline{M} = 0$	Quiotepec	Carib	Maidu (Gbeya)
	Chinantec	Cashinahua	Minnan (Hausa)
	Gbeya	Eskimo	Chinese (Quechua)
	Susu	(Inupik)	Margi
		Galla	Sinhalese
		(Hausa)	Songhai
		Kayan	(Susu)
		Kurukh	(Tagalog)
			(Telefol)

Figure 5b. Classification of languages by syllable structure.

$\bar{I} = 1, \quad \bar{F} = 1$

	$\bar{M} = 1$	$\bar{M} = 2$	$\bar{M} = 3$
$\underline{I} = 1, \underline{M} = 1$		(Egyptian Arabic)	(Bashkir) Zoque
$\underline{I} = 1, \underline{M} = 0$		Arapaho Baure (Eg. Arabic)	Koya Turkish (Zoque)
$\underline{I} = 0, \underline{M} = 1$		Bushong Eg. Arabic Njungar	Karačay Oykangand
$\underline{I} = 0, \underline{M} = 0$		Yukaghir	Chipaya Kolami Quechua

Figure 5c. Classification of languages by syllable structure.
$\bar{I} = 1, \bar{F} = 2$

	$\bar{M} = 1$	$\bar{M} = 2$	$\bar{M} = 3$
$\underline{I} = 1, \underline{M} = 1$		Acoma Mazateco	
$\underline{I} = 1, \underline{M} = 0$		Huichol	
$\underline{I} = 0, \underline{M} = 1$		Itonama Osage	Mazahua (Kannada)
$\underline{I} = 0, \underline{M} = 0$		Ijo	

Figure 5d. Classification of languages by syllable structure.
$\bar{I} = 2, \bar{F} = 0$

	$\bar{M} = 1$	$\bar{M} = 2$	$\bar{M} = 3$
$\underline{I} = 1$, $\underline{M} = 1$		Amuzgo	Halang
		Cham	Pacoh
		Kaingáng	Sedang
		Khmer	Thai
		Western Desert (Njamumada)	
$\underline{I} = 1$, $\underline{M} = 0$			Khmu
$\underline{I} = 0$, $\underline{M} = 1$		Alabaman	Jeh
		Dan	Kashmiri
		Tamazight	Palaung
		(Thai)	Yakur
$\underline{I} = 0$, $\underline{M} = 0$		Usila Chinantec	Apinayé
		Chukchee	Bengali
		Garo	Eskimo (Yupik)
		Japanese	(Luo)
		Koryak	Madurese
		Lendu	Spanish
		Middle Wahgi	
		Sierra Nahuat	
		Tetelcingo Nahuatl	
		Sara	
		Telefol	
		Trique	

Figure 5e. Classification of languages by syllable structure, $\bar{I} = 2$, $\bar{F} = 1$

	$\bar{M} = 1$	$\bar{M} = 2$	$\bar{M} = 3$	
$\underline{I} = 1, \underline{M} = 1$			Aguacatec	
			Bashkir	
			Totontepec Mixe	
			Squamish	
			Wiyot	
			(Zoque)	
$\underline{I} = 1, \underline{M} = 0$			(Turkish)	
			(Zoque)	
$\underline{I} = 0, \underline{M} = 1$		Pame	Czech	
			(Karačay)	
			Palestinian Arabic	
			Shilha	
$\underline{I} = 0, \underline{M} = 0$		Akan	Amharic	Irish
			(Chipaya)	Latvian
			(Czech)	Luo
			English	Norwegian
			Estonian	Pashto
			Georgian	Russian
			Gujarati	Totonaco
			Hungarian	Walapai

Figure 5f. Classification of languages by syllable structure.
$\bar{I} = 2, \bar{F} = 2$

2.3 More observations. The most striking feature of the classifications is that there are no languages possessing initial clusters ($\bar{I} = 2$) or final clusters ($\bar{F} = 2$) which lack medial clusters ($\bar{M} = 0$):

(5) If a language has initial or final clusters, then it also has medial clusters.

The preference for medial clusters has been noted before. Trnka (1936) stated that

> The occurrence of consonantal combinations at the beginning, or end, of words involves the possibility of their use between vowels, but not vice-versa (cf. pl, kl, kw as against mp, mb, lp, lb in English).

This claim is much stronger than (5), of course. In fact, it is too strong, and counterexamples are not too hard to find. CARIB, TELEFOL, and IJO all have recently formed initial clusters in C + r, but possess no medial C + r clusters. It is curious that Trubetzkoy (1969, 247) restates Trnka's rule, without explanation, in a more conservative form equivalent to (5):

> Trnka's next rule states: Languages that permit combination of consonants initially or finally also permit consonantal combinations medially. This law really seems to be valid for languages with polysyllabic words.

To all appearances, this regularity is not an isolated fact, but rather an extreme manifestation of a general harmony among initial, medial, and final clusters. Not only are languages with terminal, but no medial, clusters disfavored, but the opposite imbalance is rarely found. For example, KANNADA is the only language in the sample with medial clusters of three segments with no initial clusters and no final consonants. The proportion of languages with three-segment medial clusters that have initial or final clusters is greater than for those with simpler medial clusters. The pattern of positive correlation between the clustering parameters is strong and consistent. It surely represents some linguistically significant regularity.

Some observations about particular syllable inventories are of interest. No language with a single syllable type was found. However, one language, PIRO, an Arawakan language of Peru, comes very close (Matteson and Pike 1958, Matteson 1965). PIRO has undergone extensive syllabic organization. Thus it has frequent syllabic consonants and sequences of consonants separated by predictable vowel transitions. Where consonants are identical, full loss has occurred, resulting originally in geminate consonants. These are still retained in a neighboring dialect. In PIRO they have undergone simplification so that today the only long marginal segments are m, n, w, and y, and these vary freely with the corresponding short segments preceded by a long vowel. (The long nasals apparently are at least sometimes heterosyllabic.) A further development of interest is the reduction of vowels in syllables with glide onsets (wV, yV), giving reflexes transcribed as syllabic glides [$\underset{\text{i}}{\text{w}}$, $\underset{\text{i}}{\text{y}}$]. Since these are apparently not fricatives ([β] is given as a variant occurring before front vowels), these segments must be taken to be vowels.

The syllable types of PIRO are then principally CV (and $\underset{\text{i}}{\text{C}}$), with marginal occurrences of CVC and V syllables. The closed syllables have virtually disappeared, but, it appears, before this could happen, V syllables arose. (Most other closely related languages have initial vowels. Compare IPURINÃ ikiwi, PIRO xiwu, 'head'; MACHIGUENGA ociti, PIRO kiti, 'dog'; IPURINÃ amu-či, CAMPA abu-či, PIRO hapu 'road' (Noble 1965).)

The sample also includes one instance of a language for which the implication VC → V does not hold. NORTH PALESTINIAN ARABIC has initial vowels only phrase-initially before consonant clusters, and possesses no vowel sequences (Blanc 1953).

The distribution of languages with respect to the minimum margin parameters M (triples) and I (initial vowels) is presented in Figure 3.

It shows that 28 languages of the sample (those with $\underline{M} = \underline{I} = 1$) lack syllables without onsets. From an inspection of Figure 2, it appears that the distribution of languages with respect to these parameters is reasonably independent of the values for $\bar{I}$, $\bar{M}$, and $\bar{F}$. The interesting point about the observed distribution is that both hiatus and initial vowels occur frequently, although their typically restricted patterns of occurrence in individual languages points to a high degree of markedness.

2.4 Some characteristics of subtypes. Eventual understanding of the nature of syllable structure will require the careful study of regularities within the types defined for the purposes of this study. Although no attempt has been made to explore the following topics in depth, some interesting observations can be made on the basis of our data.

2.4.1 Syllable division. There is one kind of language with uniform syllable division, in which medial margins are always assigned to nuclei in the same way, and another in which division is not uniform. (I abstract away from the variation in distinctness of margin adherence among languages.) The differences in division are attributable to one or more of four factors: (1) inherent phonetic differences of clusters (SPANISH si.glo, al.to), (2) the presence of a constituent boundary (SPANISH a.brigo, sub.rayar), (3) effect of accent placement (ENGLISH 'dis.tal, di.'still), and possibly (4) pattern of occurrence of components of the cluster in initial and final positions (YAKUR ɛ̀.blá, yàm.lè stops not occurring finally, bl- ocurring initially, although factor (1) is also a possible explanation). The fourth factor is obviously closely related to factor (2).

The reader is cautioned that the data for this section are sparse, since many descriptions do not discuss syllable division, often perhaps because division is indeterminate for the language. The generalizations should be evaluated accordingly.

With a single marginal segment there are two possible divisions, V.CV and VC.V (ignoring ambisyllabic margins). The latter is highly marked:

(6) Languages with uniform syllabic division do not have VC.V division.

VC.V division does not usually occur without a concomitant constituent boundary. The effect of preceding accent appears rather to produce an ambisyllabic margin or a geminate, as in MADURESE (Stevens 1968), or as some postulate for ESKIMO (Ulving 1953). Intervocalic glottal stop, frequently indeterminate, may sometimes be perceived as adhering to the preceding vowel because it occurs only finally elsewhere. The following generalization is ventured:

(7) Languages without initial vowels or final consonants do not have VC.V division.[4]

Two marginal segments may be divided in three ways, VC.CV, V.CCV, and VCC.V.

(8) No languages have only VCC.V division of two-segment clusters.

---

[4] The reader will note the connection with distributional theories of syllabification, which are based on the concept that the initial and final clusters of medial syllables conform to the same constraints as initial and final syllables: Kuryłowicz (1948), O'Connor and Trim (1953), Haugen (1956), and Pulgram (1965, 1969). A restatement of the concept as an empirical regularity might be called the Word-terminal condition:

> If an intersyllabic sequence is divisible into occurring word-initial and word-final clusters, then the perceived boundary does not fall between non-occurring clusters.

This condition has astonishing generality. I know of only one exception, HUICHOL. For a discussion of distributional syllabification, see Bell (1970b), Section 2.1.4.3.

Note that (7) is not a corollary of the Word-terminal condition, since the latter does not apply to consonants occurring only medially.

PAGES 45 + 46 WERE

[redacted] MISSING FROM THE DOCUMENT THAT WAS SUBMITTED TO ERIC DOCUMENT REPRODUCTION SERVICE.

they are dissolved are important questions. The nature of word formation is obviously central to the problem. This is interesting, because we discover later that word formation is important in other ways to the understanding of the dynamics of syllable structure.

2.4.2 Margins. The parametric typology evidently conceals much of the structure of syllable margins by distinguishing them only by the number of their segments. Greenberg (1965) has investigated certain regularities of initial and final clusters. The aspect of margin subtypes most closely connected with the present study is the nature of 'limit' subtypes -- the first clusters or consonants to be acquired, the last to be lost.

Among initial clusters, there is a marked preference for consonant + liquid clusters. Out of 32 languages with either recently formed initial clusters or with very restricted cluster types, 21 had only clusters of the form consonant + liquid.[5] In 9 languages the clusters were recently formed: AKAN (br), CARIB (C+r)[6], GREBO (labial and velar C+l), IJO (voiced P+r)[7], MIDDLE WAHGI (C+r), SARA (mr, ^{m}br, gr), TELEFOL (fr, br, dr, tr, kr), the NJANGUMADA dialect of WESTERN DESERT (br, gr), and YAKUR (pl, bl). For 12 languages there was no information about the origin of the clusters: DAN (C+l), HALANG (C+r, C+l), JEH (P+r, l; m+r, l; s+r, l), KAINGÁNG (pr, mr, kr, ŋr), KASHMIRI (br, tr), KHMU (P+r, l; sr), LUSHAI (t+r, l; t^{h}+r, l), PACOH (P+r, l; but not tl), PALAUNG (labial and velar P+r, l), SEDANG (pl, ^{m}pl, ml, kl, ^{ŋ}kl, mr, ŋr), and THAI (P+r, l, P^{h}+r, l).

---

[5] Some of the languages also have forms in C+w and C+y, which are treated here as single segments, according to the assumptions of Section A.3.2.

[6] The abbreviations C = consonant, P = stop, F = fricative, L = liquid, and N = nasal are used.

[7] Williamson (1967) claims that these forms are dialect borrowings.

The languages of South-East Asia inflate the count for C+L clusters, but the evidence is strong even so. Within the group there is a striking preference for labials and velars to combine with liquids. Only LUSHAI has no labials or velars among its C+L sequences.[8] In nine of the other languages, alveolars are absent.

The development of C+L clusters in DAN and GREBO is interesting. They appear to be reflexes of initial CVd sequences in DAN (Bearth and Zemp 1967), and of CVdV (giving CIV) and CVnV (giving CĨV) in GREBO (Innes 1966).

The only other cluster type to emerge alone is fricative+consonant, found nascent in four languages: ACOMA (s, š+P), ITONAMA (s+C), SIERRA NAHAUT (s, š+C), and TETELCINGO NAHUATL (x+C).

All but one of the other instances involve at least one of these cluster types plus others. MADURESE is supporting evidence for the preferred status of C+L clusters. Its clusters result from a recent vowel loss that has been completed in Western dialects, but is still optional in Eastern dialects. There the syncopated forms occur in decreasing frequency in the following order: C+r, l; C+N; C+j, w; s+P (Stevens 1968). Other examples are BASHKIR (P+r, l and at least ks), CHUKCHEE (s+C and r, l+C), JAPANESE (F+P, S+N, ks, ks), KHMER (P+r, s+C, P+s) and LUO (k+l, r; sm; ks; m+l, r, s, č). HUICHOL is the only exception. CV morphemes in p, m, n, s syncopate in some contexts, producing CC sequences. No initial clusters in nC- or sC- are attested.

High vowel loss is the process of formation in the languages developing just F+C clusters; both reduced-vowel loss and high-vowel loss are found in the others.

---

[8] The LUSHAI sequences tr, t^hr have been variously treated as retroflex stops (Burling 1957), and as affricates (Henderson 1948). Burling's description of the l in tl and t^hl as "only vaguely voiced" suggest that they too may be affricates.

The final consonants of 39 languages and dialects of the sample are listed in Figures 6 and 7. The languages were chosen because their final margins were restricted to a few types or a few members of types.

Language	N	L	P	F
Carib	m ŋ			
Dan	n ŋ			
Dogon (Kamba)	m ɳ ŋ			
Dogon (Toro)	Vm ~ Ṽ			
Sentani	m ~ ^{d}m ŋ			
Lendu		r		
Burmese			ʔ	
Campa			ʔ	
Quiotepec Chinantec			ʔ	
Cofan			ʔ	
Kaingáng			ʔ	
Lahu			ʔ	
Trique			ʔ	
Isthmus Zapotec			ʔ	
Cashinahua				s s̃ x
Japanese				s

Figure 6. Languages with restricted final margins of a single type.

Language	N	L	P	F
Gugu-Yalanji	n $\overset{d}{\sim}$ n ɲ	ɾ r l		
Mandarin	n ŋ	l		
Margi	m n ŋ	r l		
Songhai (xl)[a]	m	r		
Wik-Munkan	m n ŋ	r		
Yahad	m n ŋ	r		
Yakur	m n ŋ	l		
Cantonese	m n ŋ		p t k	
Lalana Chinantec	n N		ʔ	
Hakka	m n ŋ		p t k	
Maru	m n ŋ		p t k ʔ	
Yay	m n ŋ		p t k	
Hausa (xl)[a]	m n			f
Usila Chinantec			ʔ	ɣ
Guarani			ʔ	x ~ h
Hottentot (xl)[a]			p ts	s
Akan	m n	l	ʔ	
Fulani	m n	l r	d ɗ	
Gbeya	m n ŋ	l r	p t k	
Berti	ŋ	r		s
Cayapa	ŋ		ʔ	s h
Cora	n		ʔ	s h

[a](xl) = (excluding loans)

Figure 7. Languages with restricted final margins of multiple types.

An obvious generalization from this data is that

(II) If a language's only final margins are stops, then they are glottal stops.[9]

Final glottal stop is the relic of a previous system of final stops in many cases. The process in progress in CHAM, where only p of b, p, t, c, and k does not have a glottal stop alternant in final position. (CHAM has final nasals and liquids also.)

The data also suggests a certain preference for nasals in final position. With few exceptions, if a language has more than one type of final margin, it has final nasals. The exceptions are USILA CHINANTEC (where ɣ has developed after a and i and may be a nuclear satellite), GUARANI (where ʔ and x~h occur after stressed vowels before pause, ʔ after low vowels, x~h after high), and HOTTENTOT (whose only finals are suffixes). The special nature of the exceptions suggests that the data may represent underlying process regularities. It is worth noting that some CHINESE languages have lost final stops, but all have retained final nasals. Possible differences between systems that are reduced from previous fuller systems and those that are recent creations of vowel loss may also prove enlightening.

3. Dynamics of syllable structure

The concept that syllable structure is an organization of segment strings was the basis for the typological classification of Section 1.4. The concept is also central to the nature of the dynamics of syllable structure.

---

[9] It would not be surprising to find exceptions to this generalization. The actual situation is of considerable interest. If (II) holds in fact, one diachronic implication is that final stops are not lost without passing through an intermediate glottal stop stage. (Excepted are cases of simultaneous loss of final stops, highly unlikely I think, and possible instances of a unique suffix morph consisting of a stop.)

I say that a linguistic form changes syllable structure when the number of its segments change and/or the syllabic organization of its segment string changes. Loss of an intervocalic consonant producing hiatus is segment change without syllabic reorganization (original two syllables remain), as opposed to intervocalic consonant loss producing a long vowel (two syllables collapse to one). The vocalization of a marginal element, as in HAUSA ɓauna < *bakna, is an example of syllabic reorganization.

The usual outcome of such processes is a change of a language's subtype, in the sense of the last section. A language may lose final consonants in all forms, while contemporaneous vowel loss is operating in the opposite sense, creating new instances of final consonants. Something of the sort happened in the development of FRENCH, so that at no time did closed syllables entirely disappear. It is only when the processes culminate to eliminate an existing syllable structure or to create a new one, that, within our skeletal typology, a language is considered to have changed its state.

A basic question raised by the present approach is whether any special importance is to be attached to the culminative changes, say the loss of the very last final consonant, as opposed to previous losses, or as opposed to the apparently 'directionless' losses in a language well-endowed with final margins. Conversely, is it necessary to assume that processes creating clusters for the first time differ from those creating clusters when they already exist? The results of the investigation do not reach far enough to confront this mountain of an issue, but its presence was felt, and I hope that a route has been determined that may lead future parties within striking distance.

3.1 Processes that shape syllable structure. The phonetic processes that shape syllable structure are vowel loss, nuclear fusion,

cluster simplification, consonant[10] loss, consonant vocalization, vowel epenthesis, consonant epenthesis, and metathesis. Morphological processes are word combination, analogic extension of inflectional and derivational elements, and metanalysis.

3.1.1 Major phonetic processes. Vowel loss may be conditioned by at least the following elements:

Quality of the vowel: High vowels and reduced vowels are most prone to loss. However, loss does not appear to be limited to a small subset of vowels, as is the case for consonants. For example, in JAPANESE, the high vowels are most frequently lost; e, o, and a are also lost, but practically only between identical unvoiced stops (Wenck 1954).

Other phonetic features. Accented vowels (whether by length, pitch or stress) resist loss. Loss is sometimes favored in pre- and post-stress positions.

Nuclear structure. Complex nuclei probably cannot be lost directly. It is for this reason that I speak of vowel loss, rather than the more general category of nucleus loss.

Position with respect to word and phrase boundary. This factor is dicussed in Section 3.2.2.3.

Constituent membership. Vowel loss is sometimes restricted to grammatical formatives.

Segmental context. Important as this factor is, it is not fully understood. Vowels next to vowels are unstable. Otherwise particular marginal contexts may be favored. Voiceless consonants, especially sibilants, are a favorite context for devoiced vowel loss, but the conditioning is probably crucial for the devoicing, not the loss. Loss next

---

[10] The term 'marginal loss' being useless in the sense of loss of marginal elements, I will use the term consonantal instead. Loss of marginal glides is thus included.

to sonants is common, found in RUSSIAN (pomolodel [p mɨ dʲél] Avanesov 1968) and in many of the examples of nascent clusters in #Cr given in Section 2.4.3. Between identical consonants is another favored environment: ELLICE puaka 'pig' te puaka 'the pig', tusi 'letter', t^{h}usi 'the letter' (Milner 1958) and MAORI ki róto ti te wāi ~ ki rót te wāi 'into the water' (Holmer 1966). A neighboring vowel can be a favoring circumstance, as indicated by CARIB fast speech alternants of #CV$_1$rV$_1$X in #CrV$_1$X (Hoff 1968).

The effect of vowel loss on the syllable structure depends on what syllabic reorganization takes place. Take initial-syllable vowel loss as an example. The string #CVCV becomes #CCV. With syllabic reorganization, an initial cluster is obtained (unless, as perhaps in the example of ELLICE above, there is simultaneous collapse of the two marginal segments into one complex one). If there is no reorgarization, two syllables remain, and #CCV is the result.[11]

Nuclear fusion, the collapse of a sequence of two nuclei into one, is the result of several processes which I find best to consider as varieties of syllabic reorganization. The single nucleus that is formed may be long or short, have the quality of one of the original vowels or one in between, or maintain both qualities as a diphthong. The possible outcomes (excluding length) are illustrated by SUSU combinations of the suffix i with vowel-final nouns: if the noun's vowel is front (i e ɛ) the

---

[11] Syllabic reorganization associated with medial vowel loss is sometimes a source of non-uniform syllable division. In some GERMAN dialects, verbs of the form like radelten, segelten lose their second vowel. The variant outcomes of the loss are found, [rat.lən] - [ra.dlən] and [zek.lən] - [ze.glən], corresponding to which syllable the margin of the lost syllable adhers to (Burckhard Mohr 1970). Medial -CC- clusters in EASTERN MADURESE are generally divided C.C, and except for some dialects, clusters of stop+liquid (PL) have a geminate stop (P.PL). But Cr and Cl clusters recently formed by vowel loss are tautosyllabic, makaluwar > ma.kluwar.

new nucleus is the front vowel ('elision' of i); if the vowel is a, the new nucleus is ɛ ('contraction' of a+i); if the vowel is back (u o ɔ), the new nucleus is a diphthong with front offglide (u̯i, o̯e, ɔ̯ɛ) ('devocalization' of i), alternating with a front vowel (i e ɛ) (Houis 1963).

The processes of cluster simplification and consonant loss vary greatly with the features of the cluster or consonant. The segments of a cluster may be reduced by contraction as with the MARGI alveodentals p̂t, etc., or TRONDHEIM NORWEGIAN [ɭ] from rl; loss of a segment as in KASHMIRI mb# > m#, indigenous st# > s# (Morgenstierne 1941), or YAKUR 1+C > C as in yɔ̀l 'snake' + tú 'head' yɔ̀tú (Bendor-Samuel 1969); or vocalization of a segment as in HAUSA ɓauna < *bakna 'buffalo' (Newman and Ma 1966). FRENCH autre < LATIN alter.

In contrast to vowels, it does not appear that all consonants may be lost directly.[12] It is at least rare for segments other than nasals, glides, larygeals, liquids, and the voiced fricatives β, ɣ, and ɮ to be lost directly. The loss of other consonants, to my knowledge, always involves passing through other forms in a general process of lenition, such as in the common sequences s > h > ∅, k > ʔ > ∅.

I also have the impression that surrounding context does not condition consonant loss as strongly as it does vowel loss. Loss of prevocalic y, w after o, u, and i respectively in FIJIAN is one instance, however (Scott 1948).

The loss of intervocalic consonants may be accompanied by syllabic reorganizations typical of nuclear fusion, as in IGNACIANO tahina 'there is nothing', taina nuriricu 'there is nothing wrong' (Ott and Ott 1967).

The influence of position with respect to word boundary is discussed in Sections 3.2.2.4 and 3.3.

---

[12] The stability of complex nuclei might be considered an analogous feature of vowel loss. The difference is that such nuclei are lacking in many languages, but all languages have stops and fricatives.

3.1.2 Minor phonetic processes. The minor processes of epenthesis and metathesis are less frequent and occur in narrower contexts.

Vowel epenthesis is commonly found in loans, as for example TURKISH tirén 'train' (Swift 1963). But this is a process of preservation of a language's syllable structure, not alteration. Note also that what are sometimes called epenthetic or intrusive vowels from a synchronic viewpoint may, from a historical point of view, be remnants of an incompleted process of vowel loss.

The most common contexts for vowel epenthesis appear to be #__CC, #__C, and C__C#. Epenthesis before initial sibilant+consonant is found in the development of Romance and Persian, among other Indo-European languages (Grammont 1933,361). NORTH PALESTINIAN ARABIC initial clusters are preceded by i at the beginning of a phrase (Blanc 1953). Grammont also notes the rise of a vowel before initial r in BASQUE and GASCON; a similar development is reported for some BULGARIAN TURKISH dialects by Eckmann (1962). Examples of epenthesis between two final consonants can be found in ENGLISH (el m elm, etc., (Dobson 1968), CZECH (kazen< kazn, Komárek 1958), and DUTCH (kalǝf< kalf, etc., Shetter 1958). In other contexts, I assume that epenthesis is possible, but is apparently much rarer, my cursory search not having unearthed any examples.

One does not normally think of an epenthetic vowel arising next to another vowel. The closest thing to it is the process V > V, V, which could be called nuclear fission. I do not know of any instances of the completed process, although examples of indeterminacy and variations are common. The original nucleus is complex, either a diphthong or an 'interrupted nucleus' (as in DANISH nuclei with stǿd, and nuclei transcribed as V?V in many American Indian languages). In some cases, in slow or emphatic speech, the nucleus may split. An example of a diphthong alternation is GREBO kwa ~ ko.a, the latter occurring in slow speech (Innes 1966).

Grammont (1933, 232ff) cites a number of examples of consonantal epenthesis. The most common environment appears to be intervocalic. Grammont's examples are of the rise of the margins w and y in this position, to which can be added ʔ, h, and ɣ as possibilities. HOTTENTOT roots in CV.V, which have diphthongal variants in rapid speech, also are sometimes CVʔV. It is interesting that the roots in CV.N never have *CVʔN variants (Beach 1938). In KOLAMI, an epenthetic h separates the junction of two identical vowels, but nonidentical vowels remain in hiatus (Emeneau 1955). CHÁCOBO vowels combine freely to form tautosyllabic nuclei; the sequence /ia/ has a frequent disyllabic variant [iɣa] (Prost 1967), another example of nuclear fission.

Epenthesis of consonants next to consonants is also well-known. Grammont cites -tst- < -tt-, -dzd- < -dd- posited for Proto-IE, str- (LATVIAN strawe) from sr- (SANSKRIT sravati), and mbr < mr, ndr < nr from FRENCH and GREEK (e.g. FRENCH chambre < *chamre). Gemination, as in MADURESE, where stop+liquid clusters are -P.PL- medially and all consonants are geminated following /ə/ (Stevens 1968), is another example. Epenthesis in prevocalic and postvocalic position is rarer. The case of the addition of final t and k in MARU appears to be one sensational instance (Burling 1966, Lyovin 1968). Burgess and Ham (1968) report the use of an epenthetic -k after stressed open syllables in APINAYÉ as an element of emphasis. The epenthetic -k is retained if a stop or nasal follows, whereas lexical k in this position is normally lost. SENTANI final stressed /i/ is often lengthened to [ij] and sometimes even [iɣ] (Cowan 1965), and SECOYA [ĩ] alternates with [ŋ] after a nasalized peak: [haĩ] ~ [haŋ] (Johnson and Peeke 1962).

Metathesis of consonant and vowel has the potential to change syllabic structure in two dimensions. For example, the change sxala < saxla 'kid' in EASTERN ARABIAN ARABIC is an instance of a regular development (C)aG > (C)Ga, where G stands for the segments x, ɣ, ʕ,

ŋ, and h, reported by Johnstone (1967a). It simplifies medial clusters and introduces new initial clusters. More examples could be cited. It is very difficult, however, to rule out the possibility that vowel-consonant metathesis is not the result of vowel epenthesis followed by vowel loss, or in other instances, syncope and consonant syllabification followed by vowel epenthesis (e.g. CVC > CC > VCC). In any case, metathesis appears to be a negligible factor in the dynamics of syllable structure, and it is not considered further.

3.1.3 Morphological processes. TAGALOG medial clusters are limited to nasal+stop except for clusters in g+C formed by a prefix, and many others formed by reduplication (Lopez 1965). The medial clusters of THAI consist of practically any final plus any initial margin. These are syllable structures that are not explainable by the processes considered so far. The reason, obviously, is that the phonetic processes are assumed to operate over a fixed domain, loosely called the (phonological) word. In fact, this domain is not immutable. Words are part of a higher constituent structure, and those most closely associated may coalesce into single prosodic units at the phonological level.[13] This process of word combination may introduce new medial margins, and in cases like TAGALOG and THAI, it is the major source of medial margins.

Since it is not unusual to find languages whose only initial or final margins are grammatical formatives (e.g. HOTTENTOT finals), one might consider whether there is another morphological process that creates new syllable structures in initial and final position by affixation. The evidence from such languages as ITONAMA, NAHUAT, and OYKANGAND indicates that such cases are more probably the result of phonetic processes. The only initial clusters in ITONAMA and

---

[13] But not necessarily. For example, NJUNGAR reduplications retain equal stress and open juncture between constituents (Douglas 1968).

and NAHUAT have a prefix element, but they were formed recently by vowel loss. In ITONAMA personal pronoun prefixes as-, is-, us- became s- before consonants (Liccardi and Grimes 1968) and a ši-prefix was syncopated in SIERRA NAHUAT (Key and Key 1953). As we have mentioned earlier, the only initial consonants in Oykangand are found in reduced forms of a handful of words, some of which appear to be forerunners of a classifier system (Sommer 1968, Dixon 1970). It is nevertheless plausible that most inflectional and derivational elements have their ultimate origin in ancient word combinations that have undergone extensive reduction. I thus assume that in general the morphological complexity of initial and final clusters is the result of phonetic processes operating upon word combinations, and that affixation is not an independent process that shapes syllable structure.

This does not mean that inflection plays no role in the dynamics of syllable structure. Once inflectional elements are created, they have a life of their own, and may be replaced, extended to new forms, and borrowed. For example, final -r in CLASSICAL GREEK did not disappear by a phonetic process. Analogic processes converted some forms in final -r to vowel stem nouns, e.g. MODERN GREEK andros, CL.GR. anér 'man'; MOD. GR. mitéra, CL.GR. méter 'mother' (Ward 1940). Other forms in -r were simply displaced by new forms.

It is even conceivable that analogic spread of an affix could create new clusters. Say a class of vowel-final stems takes a suffix -s. The suffix could conceivably spread to consonant-final stems, creating clusters for the first time.

Like metathesis, I consider analogical extension to be of negligible weight in shaping syllable structure.

3.2 A Markovian state-process model of initial, medial, and final clusters. One negative conclusion emerges ineluctably from the confrontation of the occurring types of syllable structures and the

processes that are responsible for their formation. Deterministic exclusion principles of process application are unlikely to explain the distribution of syllabic types. For some of the types not occurring in our sample, notably languages without initial margins and languages with uniquely CV syllables, there exist languages (namely OYKANGAND and PIRO) which come so close to exemplifying the type that I have not the temerity to claim that they cannot occur. Where hierarchies of process application might be fruitful, they do not appear to hold. Here I have in mind the observation (5) that in our sample no languages with initial clusters occurred that did not also have medial clusters. It would be neat indeed if it were true that between consonants, vowels are lost in initial syllables only if they are also lost in medial syllables. But I see no reason to believe that such is the case.

'Why do these syllable structures not occur?' is the wrong question, and 'Why are they so rare?' is the proper one. In a crude way, the answer is obvious. The processes that lead to rare states must be much less likely to occur than those by which languages of rare types are transformed into commonly occurring ones.

Once launched in this direction, we find further progress not so obvious. Most types may evolve from several other types and develop into another three or four. Some processes play a role in a number of different transitions. The number and interplay of variables is bewildering. A systematic way of representing them is essential. For this reason I have used a Markov chain to model the dynamics of marginal clustering of syllable structures.

3.2.1 Markov chains.[14] A Markov chain is composed of a set of

[14] The reader should refer to Kemeny and Snell (1960) for a fuller treatment of Markov chains at an elementary level, and to Feller (1968) or Parzen (1962) or works cited therein for more advanced treatment.

states and the probabilities of transition between the states. In most applications, the chain is used to describe how objects move from state to state. (In the present application the objects are languages.) A basic property of the transition probabilities is that they depend only on the state of origin. It is sometimes said that Markov chains represent 'memoryless' processes, since the states traversed by an object in the past have no influence on its future course. (A finite state grammar, made familiar to linguists by Chomsky's Syntactic Structures, is essentially a Markov chain stripped of its probabilistic weights of transition between states.)

To take an example of the application of a Markov chain, let us consider languages with initial clusters as one state (State 1) and languages without initial clusters as another (State 0). Over some arbitrary time interval, a language may either change state (with probabilities p_0 and p_1) or stay in the same state (with probabilities $1-p_0$ and $1-p_1$). The process is diagrammed in Figure 8.

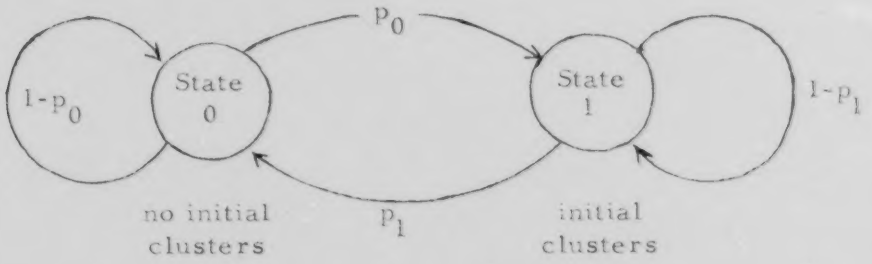


Figure 8. A two-state Markov chain representing initial clustering.

The probabilities of finding a language in a state at a given time, $X_0(t)$ and $X_1(t)$, may be expressed in terms of these transition probabilities:

$$(12)\quad X_0(t) = (1-p_0)X_0(t-1) + p_1X_1(t-1)$$
$$X_1(t) = p_0X_0(t-1) + (1-p_1)X_1(t-1)$$

For Markov chains of the kind used here,[15] for sufficiently large t, the state probabilities approach a limit which is independent of the initial state. In our example, the state probabilities are given by the solution of the equations (and the condition $X_0 + X_1 = 1$)

$$(13)\quad X_0 = (1-p_0)X_0 + p_1X_1$$

$$X_1 = p_0X_0 + (1-p_1)X_1$$

or

$$X_0 = p_1/p_0+p_1, \quad X_1 = p_0/p_0+p_1$$

Now if we had reason to believe that languages with initial clusters were twice as likely to occur as languages without, we could represent the situation by a Markov chain with state probabilities $X_0 = 1/3$ and $X_1 = 2/3$. For such a chain the probability of transition from state 0 to state 1 (p_0) would be twice that of transition in the other direction. It is this connection between state probabilities and transition probabilities that will be exploited below.

3.2.2 Construction of the model. To construct the model, many specific hypotheses about syllable structures and the processes that change them must be made. They are described in the following sections in detail. Overshadowing them are three general assumptions. First is the Markovian property. For the states postulated, the processes are independent of all past states that a language has traversed. The second, another assumption of independence, is that process likelihoods are independent of the states per se. By process I mean here a structural change plus a structural description. A state affects only whether a process's structural description is satisfied. If it is satisfied, then the

---

[15] For the Markov chains used here it is possible to reach any state from any other state, and the probability of remaining in a state is never zero. These are the properties of irreducibility and aperiodicity, and such Markov chains are called ergodic.

processes affecting final clusters will have the same effect on languages with and without initial clusters. (I recognize that this defines an idealized situation.) The third assumption is that only processes internal to a language are considered; an ideal situation in which loans do not affect syllable structure is assumed.

3.2.2.1 States. To hold the number of states in the Markov chain to a reasonable number, some classification of syllable structure coarser than the full parametric typology was required. The states defined by the parameters $\bar{I}$, $\bar{M}$, and $\bar{F}$, which represent the maximum number of marginal elements in initial, medial, and final position, were considered most suitable. Their number, 36 (3 · 4 · 3), was workable. The distribution of languages across these states exhibited interesting dependencies. The other parameters of the typology, I and M, concerning the presence of initial vowels and hiatus, appeared to be relatively independent of $\bar{I}$, $\bar{M}$, and $\bar{F}$ (Section 2.3.). They are treated separately in Section 3.3. below.

3.2.2.2 Processes. The states having been defined, it remains to establish the transition probabilities between them. The following procedure was used. First, the significant processes leading to changes of state were enumerated. These processes fell into three categories: cluster formation processes, cluster simplification processes, and word combination processes. Within each category, processes were assigned relative probabilities. No attempt was made to establish a priori the importance of the three process categories relative to each other. These relationships were expressed by constants which were adjusted in terms of the computed state probabilities in the manner described in Section 3.2.3. below. Finally, the 'rest' probabilities, the probabilities that a language does not change state, were derived from the 'nonrest' transition probabilities, since the rest and nonrest transitions must sum to

unity. It is important to keep in mind that no attempt was made to establish absolute numerical values for the transition probabilities. The numerical values that appear in the model represent the likelihoods of transition relative to each other. (This is possible because we are guaranteed, by the nature of the type of Markov chain used, that the equilibrium state probabilities are uniquely determined by the relative values of the nonrest transition probabilities.[16] I also emphasize, not

---

[16] For an ergodic Markov chain with transition probabilities p_{ij}, the stationary state probabilities x_i are given by the equations

$$
\begin{aligned}
& x_1 p_{11} + x_2 p_{21} + \cdots + x_n p_{n1} = x_1 \\
& x_1 p_{12} + x_2 p_{22} + \cdots + x_n p_{n2} = x_2 \\
& \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\quad \vdots \\
& x_1 p_{1n} + x_2 p_{2n} + \cdots + x_n p_{nn} + x_n
\end{aligned} \tag{1}
$$

or

$$
\begin{aligned}
& x_1(p_{11}-1) + x_2 p_{21} + \cdots + x_n p_{n1} = 0 \\
& x_1 p_{12} + x_2(p_{22}-1) + \cdots + x_n p_{n2} = 0 \\
& \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\quad \vdots \\
& x_1 p_{n} + x_2 p_{2n} \quad \cdots \quad + x_n(p_{nn}-1) = 0
\end{aligned} \tag{2}
$$

Now let the nonrest transition probabilities be determined up to a constant A, so that $p_{ij} = Aq_{ij}$ for $i \neq j$, and the rest probabilities be given by $p_{ii} = 1 - A\sum_j q_{ij}$. The equations (2) then become

$$
\begin{aligned}
& -Ax_1\textstyle\sum q_{1j} + Ax_2 q_{21} + \cdots + Ax_n q_{n1} = 0 \\
& Ax_1 q_{12} - Ax_2\textstyle\sum q_{2j} + \cdots + Ax_n q_{n2} = 0 \\
& \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\quad \vdots \\
& Ax_1 q_{1n} + Ax_2 q_{2n} + \cdots - Ax_n\textstyle\sum q_{nj} = 0
\end{aligned} \tag{3}
$$

Since the solution of (3) is independent of the value of A, the stationary state probabilities x_i are determined by the relative values of the nonrest transition probabilities. (I am indebted to Arthur Albert for indicating the proof of this theorem to me.)

for the last time, that although assignment of numerical values to the likelihood of processes is necessary for the application of the model, they are to be conceived as representing ranks and orders of magnitude rather than exact values.

The processes between states are diagrammed in Figure 9, 10, 11, and 19. The states are represented by a sequence of three digits, the values of the parameters $\bar{I}$, $\bar{M}$, and $\bar{F}$. For example, the state 210 represents languages with initial clusters, medial consonants but no medial clusters, and no final consonants. For simplicity, since languages with no initial consonants ($\bar{I} = 0$) are very rare, no attempt was made to represent the processes concerning their medial and final clusters. For the phonetic processes, C represents a marginal segment; V, a syllabic nucleus; #, word boundary; -V, -C, loss of element, i.e. $V \rightarrow \phi$, $C \rightarrow \phi$; and +V, +C, epenthesis of an element, $\phi \rightarrow V$, $\phi \rightarrow C$. Figure 19 shows the possible transitions that might be caused by the coalescence of two words into one. This process affects only the formation of medial clusters.

There is a basic distinction of nature between the processes of cluster formation and cluster simplification. (Formation and loss of single consonants are included in these categories.) To change a language's type in the direction of increased clustering, all that is required is for a single instance of the new cluster to be created. A language with no final consonants ($\bar{F} = 0$) moves to a state with $\bar{F} = 1$ if a single final vowel is lost. Cluster formation processes, therefore, involve a single instance of occurrence to achieve a change in type.

In the other direction, all clusters characterizing a state must be simplified. A language with final consonants ($\bar{F} = 1$) may be subject to loss of final consonants over a long period, but it is not until the last one is lost that the overall syllabic type is changed. Cluster simplification processes must be carried out to completion to achieve a change in type.

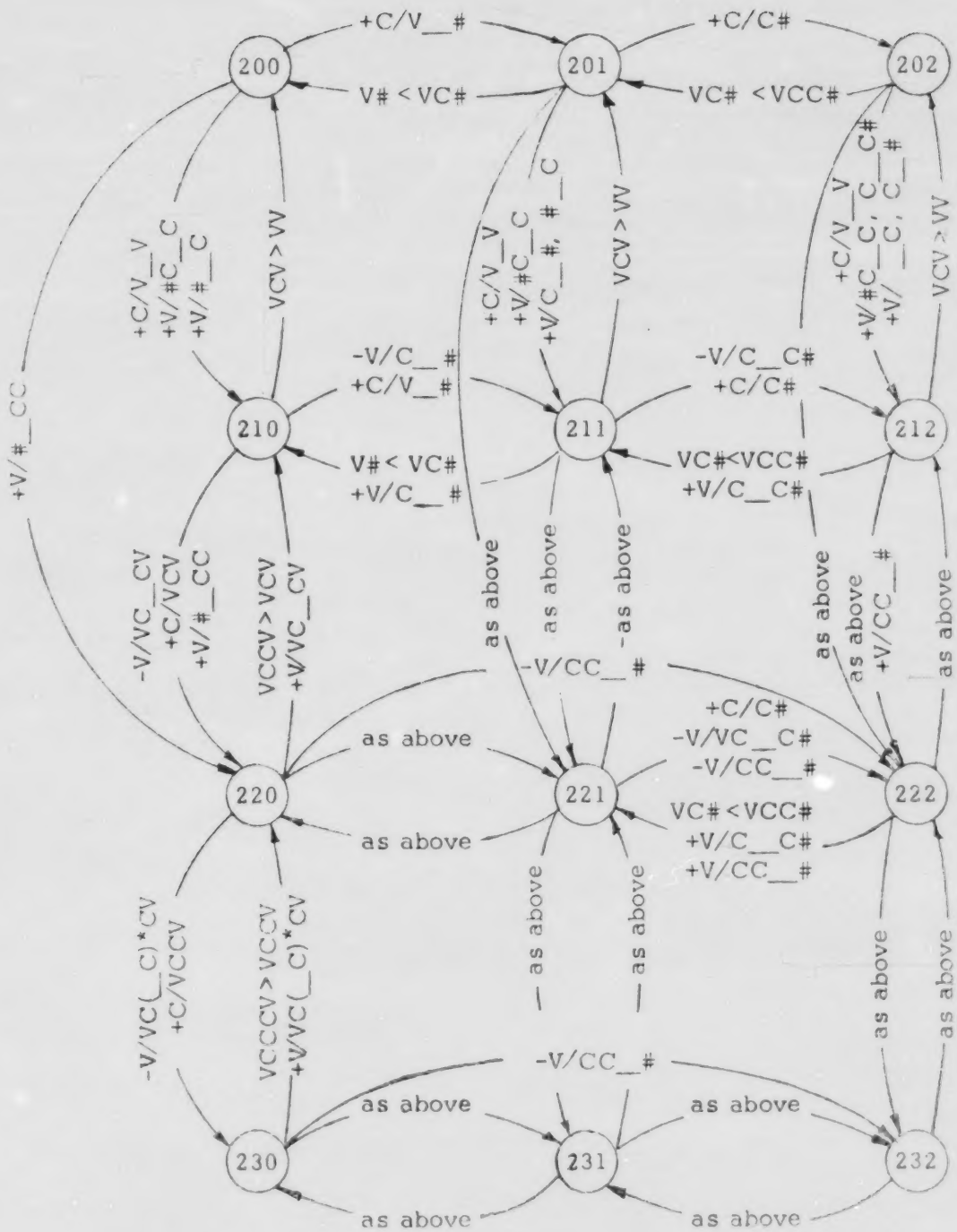


Figure 9. Processes between states with initial clusters ($\bar{I} = 2$).

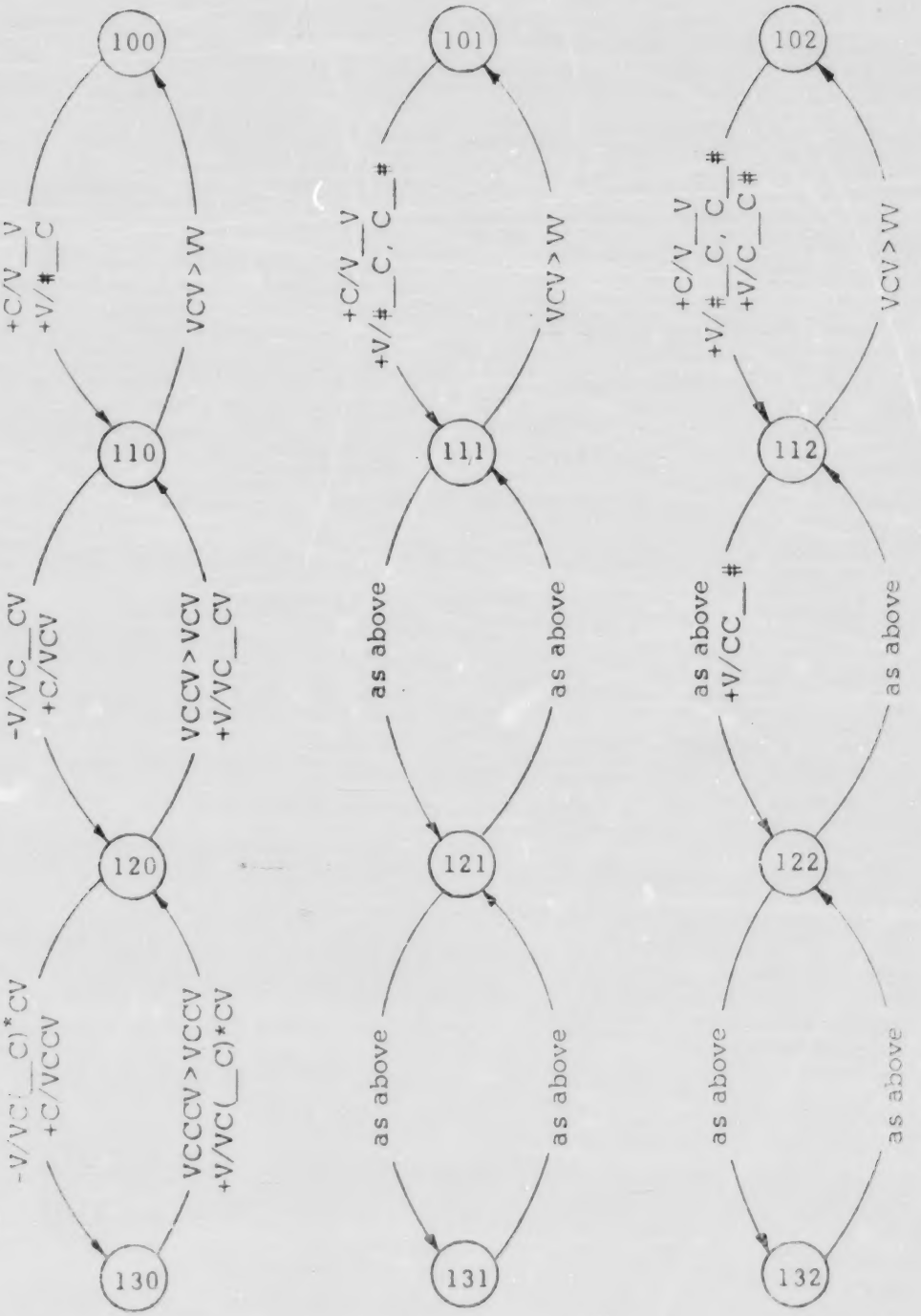


Figure 10. Processes between states with initial consonants ($\bar{I} = 1$). The processes affecting final position are the same as for $\bar{I} = 2$ (Figure 9).

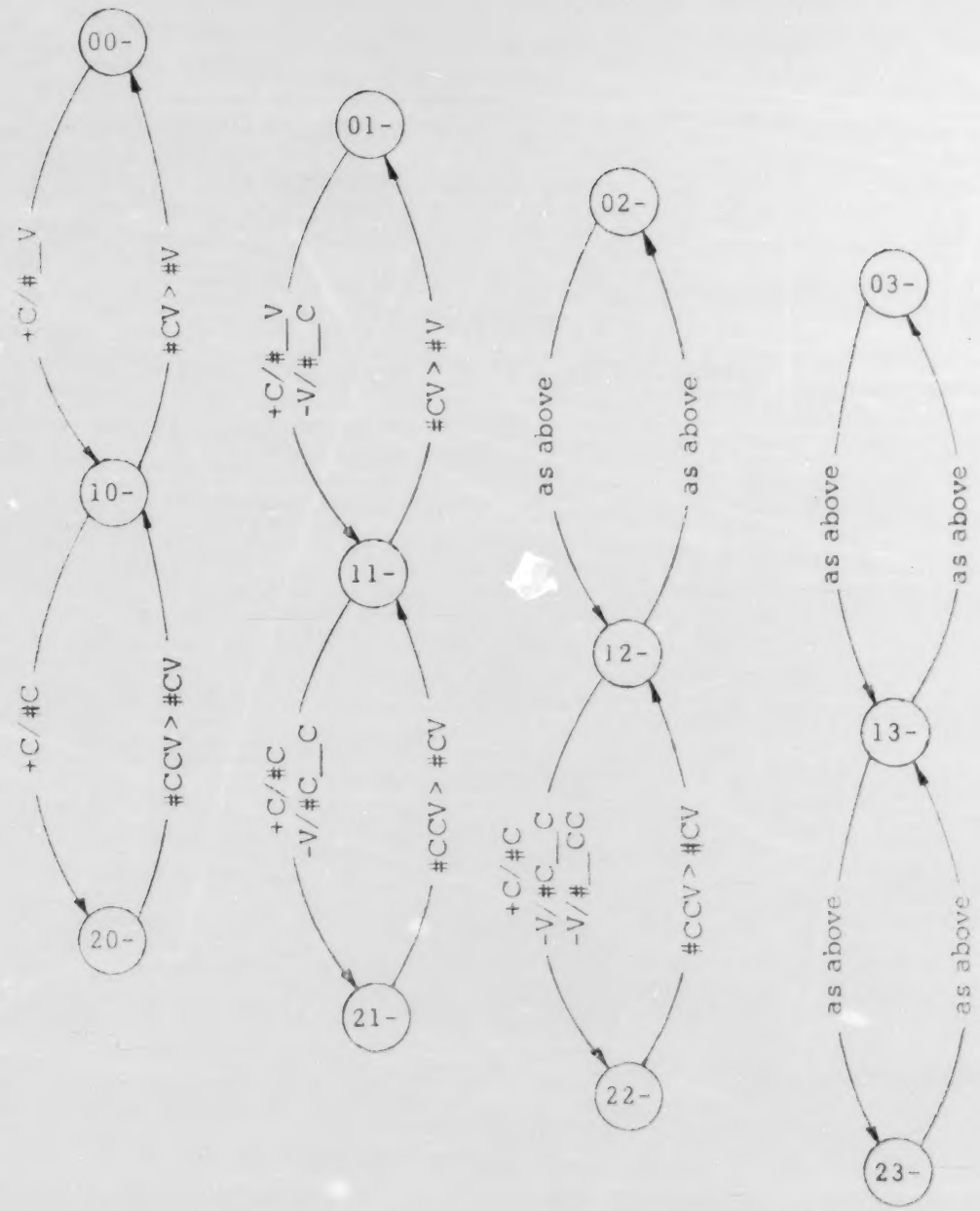


Figure 11. Processes affecting initial position. The processes are independent of the value of $\bar{F}$, hence the hyphen in place of the third digit.

3.2.2.3 Relative likelihoods of cluster formation processes. The most important phonetic process of cluster formation is vowel loss. Others of significance are consonant epenthesis and vowel epenthesis. (The latter may operate to create longer medial clusters by shifting initial or final clusters into medial position.)

Eight contexts of vowel loss are distinguished. It may occur initially, either before a single consonant or before a cluster (/#__C, /#__CC). It may occur finally, either after a single consonant or after a cluster (/C__#, /CC__#). Between consonants, vowels may be lost either in an initial syllable, in a final syllable, or in a medial syllable between two simple consonants, or in a medial syllable between a consonant and a cluster (/#C__C, /C__C#, /VC__CV, /VC(C__)*CV).

Three important factors influencing the loss of vowels are ignored here -- vowel quality, vowel length, and stress. The first is ignored because vowels of all available qualities normally occur in all positions. The 'prosodic' features of length and stress are certainly related to the language types considered here, but, for the sake of simplicity, are not treated in the present model. Within a given type all languages are lumped together, whatever their differences in accentual system or otherwise, and the weights of the processes connecting types represent averages over all languages of the type.

The effect of position within the word is the main one that must be established. Here I appeal mainly to the empirical tally of vowel loss recorded in the descriptions of languages in the sample. The 44 instances of recent vowel loss found there had the following distribution by position:

Word-initial	4
Word-medial	
First syllable	21
Medial syllable	12
Last syllable	4
Word final	14
	55

This evidence is crude, and perhaps the most that one can say for it is that it is tangible, and not just opinion. In any case, it is satisfactory for our limited needs of establishing a nonarbitrary ranking of vowel loss processes. Using it as guide, I assume the partial ordering of Figure 12.

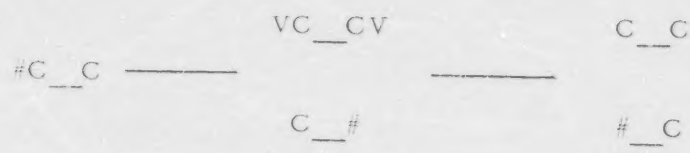


Figure 12. Partial ordering of likelihoods of vowel loss according to word position.

A second factor influencing vowel loss which must be accounted for is the consonantal environment, as was seen in Section 3.1.1. The conditioning factor may also be the number of segments in the cluster that would be produced, as in FRENCH (more or less) and presumably reflected in the phonological rule of YOKUTS, V→ Ø / VC__CV (Kisseberth 1970). I conclude that in a given position, a vowel is the more likely to be lost the fewer the marginal segments surrounding it. For example, it is assumed that medial vowel loss in the context /VC__CV is more likely than in the context /VC(C__)*CV.

This preference differs from the preferences of position in a fundamental way. In any given position, I assume that vowel loss may take place without vowel loss occurring elsewhere, but in a given position vowel loss does not take place in a more complex marginal surround unless it also takes place in a simpler one. A consequence of this may be seen in Figure 9 and 10, where from the states with medial clusters but no final vowels (120, 130, 220, 230) transition to states with final consonants (121, 131, 221, 231) may take place or a direct transition to states with final clusters (122, 132, 222, 232) may occur. On the other hand, from the states with no medial clusters (110, 210), final vowel loss can lead to only one new state (111, 211). If, therefore, loss in the context

/C__# is assigned a relative likelihood p_c and loss in the context /CC__# a lesser likelihood, say 0.5 p_c, and the two events treated independently, then a change from medial cluster states by final vowel loss will be more likely than from the states without medial clusters. I think this is inconsistent with the general assumptions of Section 3.2.2. Instead, I will assume that there is a given probability of final vowel loss, say p_f. Where final vowels are preceded both by clusters and by single consonants, the two possibilities are loss only after single consonants, and loss after both clusters and consonants. I assume further that the first is more likely, and to assign a relative weight, take it to be twice as likely.[17] The values of vowel loss transitions from the states 210, 211, 220, and 221 are illustrated in Figure 14 as an example.

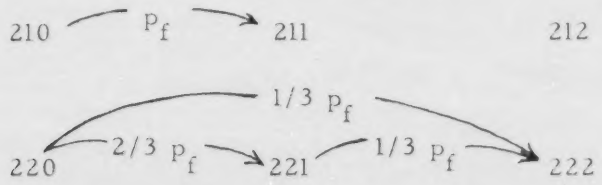


Figure 14. Illustration of relative likelihoods of transition by final vowel loss.

Initial and medial vowel loss are treated similarly. Note that there is no case where medial vowel loss may result in possible transitions to two states.

The epenthesis of consonants and vowels is less common than vowel loss. From the discussion of Section 3.1, consonant epenthesis appears to be most common in intervocalic position, and least common in the environments /#__V and /V__#; the likelihood of the sequences #C, VCV, VCCV, and C# developing an additional segment lies between

---

[17] The arbitrary choice of value is not the only source of error here; it is perfectly possible that the values for the different positions are widely different.

them. The assumed values are given in Figure 15.[18]

V__V	0.05
#C	0.01
VCV	0.01
VCCV	0.01
C#	0.01
#__V	0.003
V__#	0.003

Figure 15. Assumed relative likelihood of occurrence for consonant epenthesis.

The inclusion of vowel epenthesis among cluster formation processes doubtless appears odd. It functions in this role only by transferring initial and final clusters to medial position. Thus the environments of interest are #__C, #__CC, #C__C, C__#, CC__#, and C__C#. From the examples given in Section 3.1, epenthesis occurs in the contexts #__CC, #__C, and C__C# much more frequently than in the others. The assumed values of relative likelihood of occurrence are given in Figure 16.

#__CC	0.05
#__C	0.05
C__C#	0.05
#C__C	0.005
CC__#	0.005
C__#	0.005

Figure 16. Assumed relative likelihood of occurrence for vowel epenthesis.

---

[18] All stated values of relative likelihoods are scaled by a factor of 1/100 for actual computation. Thus the likelihood of the process $V \rightarrow \phi$ / #C__C, stated to be 0.30, would be scaled to 0.003 in the program for computing the state probabilities.

3.2.2.4 Relative likelihoods of cluster simplification processes. The most important processes of cluster simplification are consonant loss, coalescence of two consonants (as in the formation of the MARGI labio-alveolars p͡t, etc.), and the mutation of a marginal segment into a nuclear one (e.g. FRENCH autre < alter). The role of vowel epenthesis is probably negligible, since, being limited to certain kinds of clusters and segments, it is unlikely to be carried to completion. The problem of evaluating the relative likelihood of these processes is made difficult by the fact that here we are not dealing with the initiation of the process, but its culmination. More precisely, we are concerned with the attainment of simplification of all clusters or loss of all consonants in a given position, no matter whether due to the spread of a single process, say loss, or whether aided in some instances by vowel epenthesis. Therefore, in this section, when referring to a process such as #CC > #C, I mean the reduction of all such clusters in a language by whatever phonetic means.

Now the likelihood of a language losing all its final consonants, to pick one example, naturally depends upon how many final consonants the language has, what kinds they are, and how frequently their types and tokens occur. A language in which only n, s, h and ʔ occur (as in CORA) is a more likely candidate for complete loss of final syllables than one in which all consonants may appear in final position. Therefore we are dealing with but the final aspect of a complicated process taking place within each type, but completely unrepresented in our model. There is no harm in this, but eventually our assumptions about cluster simplification must accord with the evidence concerning the subtypes and the processes between them. A further consequence is that data concerning simplification of clusters is not a sufficient guide. Loss of certain consonants in intervocalic position is common. Yet languages never lose all their intervocalic consonants.

For the moment, I can only attempt to gauge the relative likelihoods indirectly. It appears to be possible to obtain a satisfactory ranking from two basic principles. The first principle states that longer clusters are weaker: the completion of the process $C^{n}_{n} > C^{n-1}_{n-1}$ for all C^{n}_{n} in a given position is more likely than the completion of the process for C^{n-1}_{n-1}. The second principle states a positional hierarchy: the process $C^{n}_{n} > C^{n-1}_{n-1}$ is less likely to be carried to completion in medial position than in a terminal position, and less likely to be completed in initial position than in final position.

In assuming the values of relative likelihood for each of the relevant cluster simplification processes, I have assumed that loss of initial and medial consonants is very much less likely than any other case, and that the others are roughly of the same order of magnitude. The values are given in Figure 17.

VCCCV > VCCCV	0.60
VCC# > VC#	0.50
#CCCV > #CV	0.40
VCCV > VCV	0.20
VC# > V#	0.20
#CV > #V	0.0004
VCV > VV	0.0002

Figure 17. Assumed relative likelihood of occurrence of completed cluster simplification processes.

3.2.2.5 Relative likelihoods of word combination. By word combination is meant the process by which two consecutive words fuse into a single one. This may or may not result in formation of medial clusters of increased length, as can be seen from Figure 19, where between several states there are no transitions attributed to word combination. This follows from the assumption that the elements of

a newly coalesced word obeyed the constraints of the language's syllable structure. A language with no initial or medial clusters and no final consonants (type 110) cannot change its syllable structure by word combination because the only possible sequences across word boundary are -V#C- and -C#V-. Loss of the word boundary creates a form already existing. On the other hand, a language of type 112 (as above, except possessing final clusters) can change to two different types by the process of word combination. Medial clusters of two consonants can be created in two says: either from a sequence -C#C- or a sequence -CC#V-. Clusters of three may be formed from the sequence -CC#C-.

The relative weights of the transitions were derived from the following assumptions.

(1) Within a given time span, the likelihood of the occurrence of word combinations is the same for all languages, whatever their syllable structure.

(2) The likelihood that a language possesses words with initial vowels is 0.7. (See Figure 4, which shows the observed ratio in our sample.)

(3) Words beginning with a vowel and words beginning with a consonant are equally likely to enter into combination, and similarly for final position. Words beginning with a cluster are 0.1 as likely to combine, and similarly for final position.

The remaining factor needed for computation of the transition probabilities is the number of compounds formed in the interval of time chosen as a unit. The simplest assumption would be that the probability that more than one compound is formed during a unit time interval is negligible. This appears unrealistic to me, since it implies a very short time interval. Language changes of the sort considered here are not instantaneous events, but require a certain amount of time for

diffusion over ideolects and lexicon before it may be said to 'happen'. I assume then that the actual situation is that in each unit interval the probability of more than one compound being independently formed is significant compared to the probability of just one. A rough approximation to the process may be made by computing the relative likelihoods of transition from each state as if 5 compounds had been formed.[19] The effect of this assumption is to shift the likelihoods in favor of forming longer clusters, as can be seen in the comparison of Figure 18.

From State 102 to State	One Combination	Five Combinations
102	0.17	0.00
112	0.48	0.14
122	0.32	0.72
132	0.03	0.14

Figure 18. Relative likelihoods of transition from State 102 by word combination.

The relative likelihoods of transitions computed according to the above assumptions are presented in Figure 19.[20]

---

[19] Strictly speaking, the transitions should be computed for each number of compounds and a weighted average of the transitions made. This nicety is hardly necessary here, but it should be remembered that the figure of 5 compounds is a nominal one and does not necessarily imply that the distribution of number of compounds formed has a mean value of 5.

[20] The details of computation are as follows: for each state, the possible combinations across word boundary are listed, and their probabilities of occurrence computed according to the assumptions about relative frequencies of cluster, consonant, and vowel terminals. From this we find the probabilities of forming clusters of 0, 1, 2, and 3 segments (p_0, p_1, p_2, p_3). The probability that out of 5 compounds, x_0 of them form 0-segment clusters; x_1, 1-segment clusters; x_2, 2-segment clusters; and x_3, 3-segment clusters is given by the multinominal

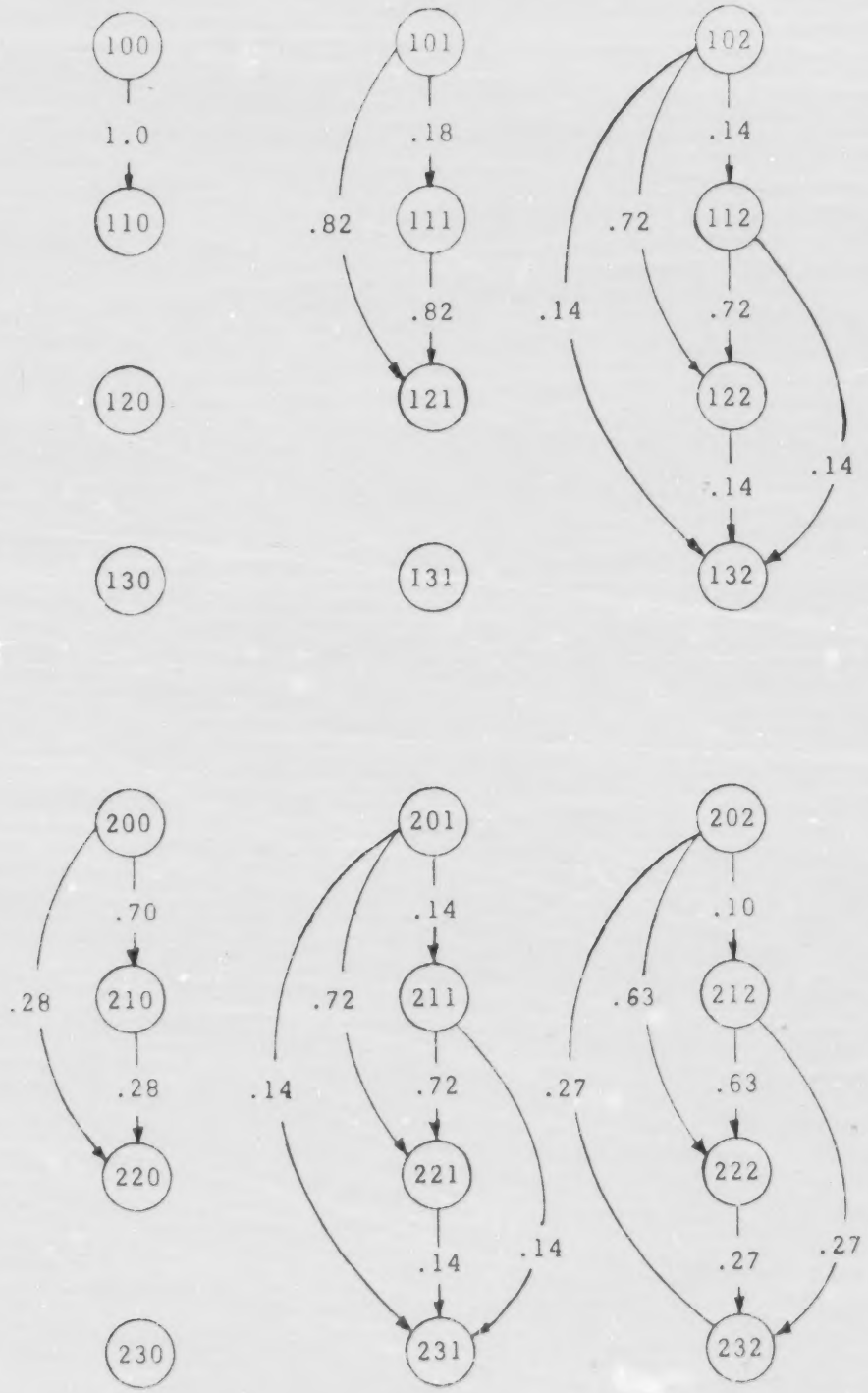


Figure 19. Relative likelihoods of word combination processes.

3.2.3 Application of the model. The components of the Markov chain model defined in the last section are a set of 36 states defined over the parameters $\bar{I}$, $\bar{M}$, and $\bar{F}$; the processes of transition between the states; and relative likelihoods of the processes, independently determined for the three process categories -- cluster formation, cluster simplification, and word combination. What is lacking is a means of relating the likelihoods of the three groups of processes. This is done by means of two constants, R and WF. R represents the relative weight of the cluster formation processes with respect to the cluster simplification processes. In the application of the model, the values of the processes given in Figures 14, 15, and 16 are multiplied by R. Similarly, values of the word combination processes (Figure 19) are multiplied by WF.

The purpose of the model is to relate the assumptions about processes described in the previous section to the observed distribution of languages over states. The distribution of Figure 3 is based

---

(Footnote 20 continued)

distribution

$$p(x_0, x_1, x_2, x_3; 5) = \frac{5!}{x_0!\, x_1!\, x_2!\, x_3!}\, p_0^{x_0} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

We seek the probabilities π_k that none of the five compounds form clusters of size greater than k and at least one forms a k-segment cluster. Using the identity

$$(a+b)^n = a^n + na^{n-1}b + \ldots + n!/k!\,(n-k)!\, a^{n-k}b^k + \ldots + b^n,$$

these probabilities may be shown to be given by the formulas

$$\pi_0 = p_0^5$$

$$\pi_k = \left(\sum_{i=0}^{k} p_i\right)^5 - \left(\sum_{i=0}^{k-1} p_i\right)^5 \quad , \quad k = 1, 2, 3.$$

The π_k's shown in Figure 19 are a weighted average of the values obtained assuming initial vowels (0.7) and the values assuming no initial vowels (0.3).

on a sample that was not originally intended to reveal relative proportions. Even though some care was taken to achieve a balanced representation of areas and families, some closely related languages were overrepresented. Accordingly, I have adjusted the distribution of Figure 3 to correct identifiable oversampling. The corrected distribution, which will be used henceforth, appears in Figure 20.[21]

	F = 0	F = 1	F = 2	
$\bar{M}$ = 1	9	5	0	
$\bar{M}$ = 2	12	31	6	$\bar{I}$ = 1
$\bar{M}$ = 3	1	2	6	
$\bar{M}$ = 1	0	0	0	
$\bar{M}$ = 2	6	20	2	$\bar{I}$ = 2
$\bar{M}$ = 3	1	13	20	

Figure 20. Corrected distribution of languages over the cluster categories.

The first question to answer is 'to what extent can the features of the synchronic distribution of syllable structures be explained by the assumptions about processes made above?'

The method of successive approximations was used to answer this question. The procedure followed these steps.

---

[21] Examining the distribution of languages by families, I concluded that the great overrepresentation of Andean-Equatorial languages (all without initial clusters) and the moderate overrepresentation of Bantu (mainly 100, where the digits stand for the values of $\bar{I}$, $\bar{M}$, and $\bar{F}$, respectively) Mon-Khmer (mainly 231) Penutian (132, 232), and Germanic (232) languages probably biased the sample. The following corrections to the observed counts were then made: 110, -1 Bantu and -1 Andean-Equatorial; 120, -2 And.-Equ.; 121, -1 And.-Eq.; 132, -1 And.-Eq. and -1 Penutian; 231, -1 Mon-Khmer; and 232, -1 Penutian and -1 Germanic.

1) Obtain the probability transitions between each state by summing the relative likelihoods of the processes between the states.[22]

2) Assign values for the constants R and WF.

3) Compute the state probabilities for the Markov chain defined by 1) and 2).[23]

4) Compare the computed state probabilities with the observed distribution of languages.

5) If necessary, modify the constants R and WF and/or the process probabilities and repeat the process from step 1) or 2).

In the first trials, WF was set to zero, and only the constant R was adjusted. The best value for R appeared to be near 1.7. For R = 1.7 and WF = 0. the expected values for a total sample of 144 languages are presented in Figure 21. (The observed number of languages is in parentheses, repeated from Figure 20.)

The results indicate that some assumptions about the processes should be modified. First, the model accords too much weight to the formation of initial clusters. The distribution with respect to final clusters is reasonably good overall, but does not reflect the asymmetries of structure found in the observed distribution. With respect to medial clusters, there appears to be an excess of languages without medial clusters and lack of those with clusters of three or more elements.

---

[22] The summation of the likelihoods is an approximation. The probability of occurrence of at least one of two events A and B is Prob(A B) = Prob(A) + Prob(B) - Prob(A B), and similar formulas hold for greater numbers of events (Feller 1968, 98). The likelihoods of the events here, however, are small enough that the second and higher order terms are negligible.

[23] The computation was done on the IBM 360/67 at the Stanford Computation Center. The method was solution of the simultaneous equations (3) of note 10 (plus the constraint $\Sigma x_i = 1$). IBM's Fortran subroutine SIMQ, converted to double precision, was used. Computation time for each solution, including compilation of the program by the WATFOR compiler, was approximately 5 seconds.

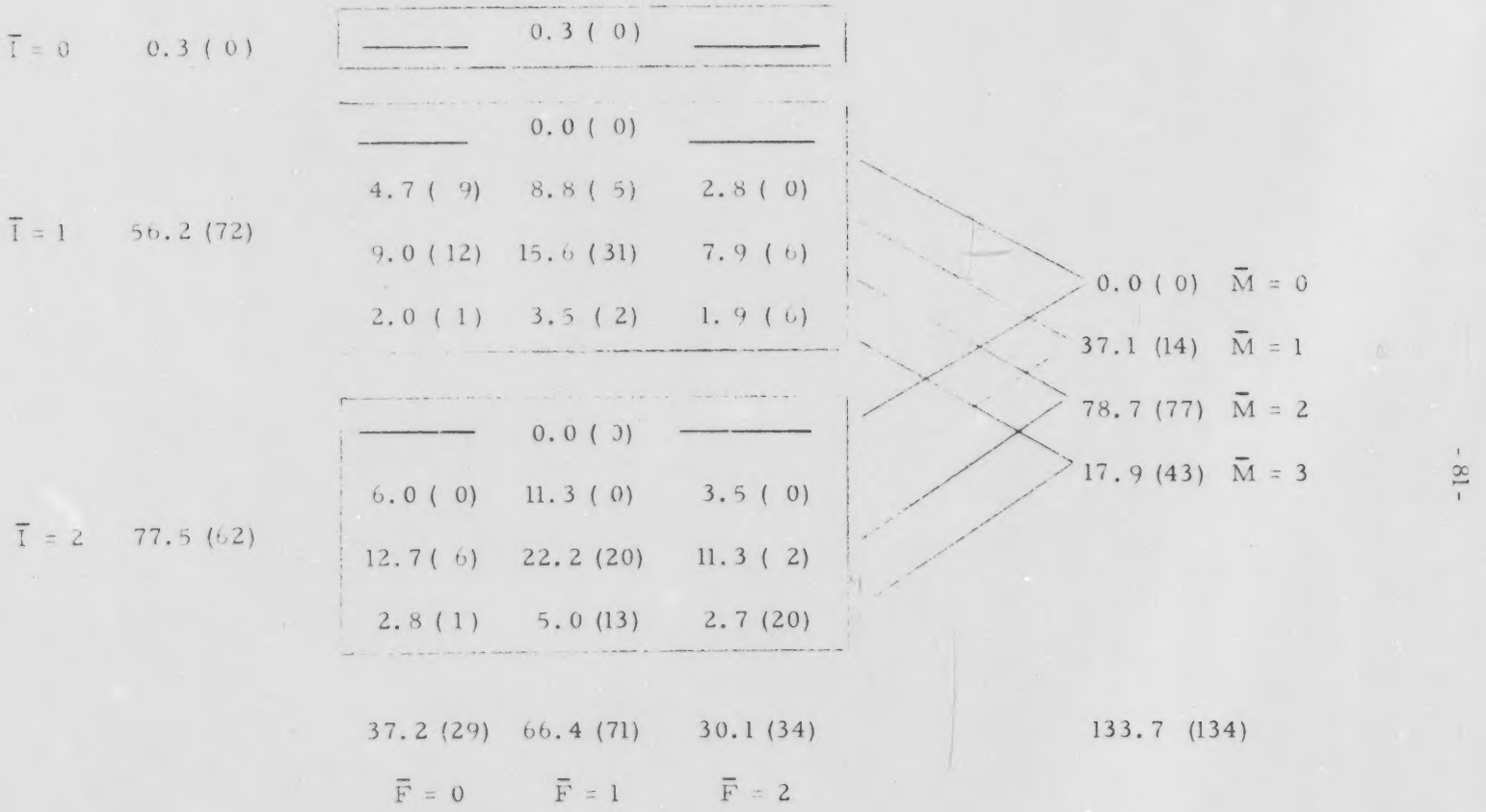


Figure 21. Computed and observed distribution of languages R = 1.7 and WF = 0.

The necessary adjustment of the relative likelihoods of the main processes affecting initial clusters (reduction of V→ ϕ/#C__C, augmentation of #CC>#C) can be done without affecting the general assumptions we have made.

The problems in the other dimensions are of a different order. Note that while overall values for $\bar{M} = 1$ are too large and those for $\bar{M} = 3$ are too small, they are approximately correct for languages with no initial clusters. This means that we want to adjust the relative likelihoods for medial cluster formation and simplification only for languages with initial clusters. However, as a comparison between Figures 9 and 10 will show, the only significant difference in the phonetic processes is that there is an additional process of epenthesis (ϕ → V/#__CC) forming medial clusters from the states 210, 211, and 212. It is absurd that this relatively rare process could explain the observed differences. Similarly, adjustment of the processes affecting final clusters will not make any qualitative differences, since the relevant processes are the same for languages with and without initial clusters.

All this, of course, is a buildup for the introduction of the word combination processes. Turning back to Figure 19 we see that they exhibit the very sort of asymmetry demanded by the data and conspicuously lacking in the structure of the phonetic processes -- a favoring of the formation of medial clusters in languages with initial or final clusters.

The indicated action is to adjust the phonetic process likelihoods to disfavor medial clustering. The new values for vowel loss and cluster simplification are given in Figure 22, with former values in parentheses if different.

Vowel Loss		Cluster Simplification	
#C__C	0.20 (0.30)	VCCCV > VCCV	0.90 (0.80)
VC__CV	0.18 (0.21)	VCC# > VC#	0.55 (0.50)
(C)C__#	0.18 (0.21)	#CCV > #CV	0.45 (0.40)
C__C#	0.09 (0.06)	VCCV > VCV	0.25 (0.20)
#__C(C)	0.06	VC# > V#	0.17 (0.20)
		#CV > #V	0.0004
		VCV > VV	0.0002

Figure 22. Adjusted values of relative likelihood for vowel loss and cluster simplification.

With the values of Figure 22 and word combination processes weighted by a factor WF = 4.0, the model yields the expected values for 134 languages whosn in Figure 23.

The result is a much better fit to the observations than the previous attempt.[24]

The marginal values (expected numbers of languages over each of the three parameters taken separately) are in quite good agreement. The significance of this has to be measured against the degree of constraint imposed upon the process likelihoods. On one hand, claims of significance must be modest because there was considerable latitude in assigning relative likelihoods according to the process rankings. After all, if there are very many sets of relative likelihoods of the processes that would produce equivalent results, it is no feat to pick one of them. On the other hand, I feel that the agreement in the marginal values does tend to substantiate the rankings of vowel and consonant loss, since there are nontrivial process rankings that the model would reject.

---

[24] Not surprisingly, I did not arrive at this result in two attempts. Presenting only two steps is a descriptive artifice. It is a realistic picture, however, abstracted away from minor bungles and hesitations caused by unfamiliarity with the sensitivity of the model. About 20 adjustments and computations were made in all.

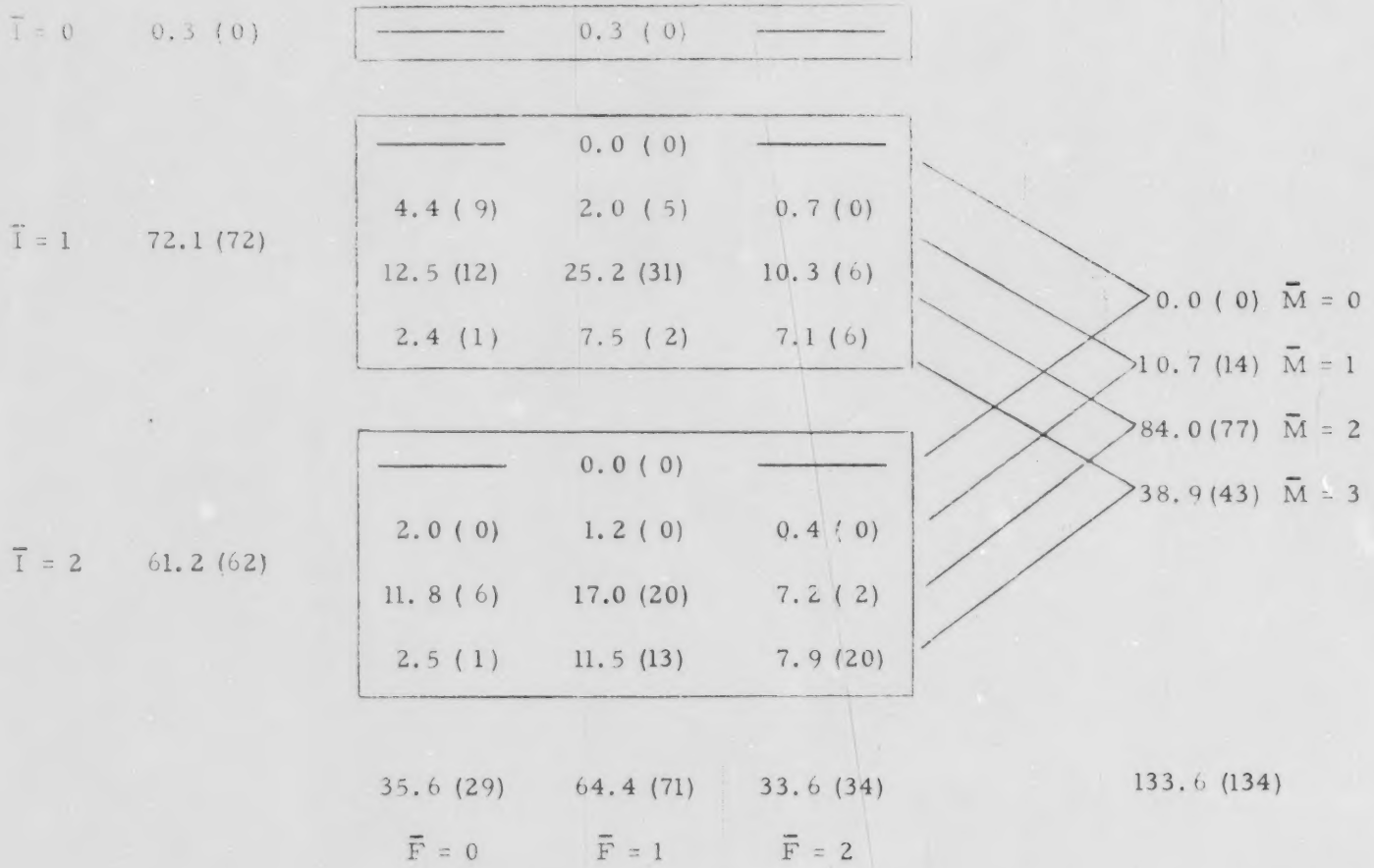


Figure 23. Computed and observed distribution of languages. R = 1.7 and WF = 4.0.

An example may make my position clear. Take first the balance between initial consonants and initial clusters. This is mainly determined by the ratio of the relative likelihoods of the process V → ϕ/#C__C (rel. likelihood 0.20) and #CC > #C (0.45). In other words, one could have obtained equally good results by raising both values or lowering them. However, each process belongs to an independent ordering. This means that if in fact one of the processes is incorrectly ordered, then the other one may well be also.

Suppose that we had assumed that #CCV > #CV was ranked lower than VCCV > VCV. (The reader will find it a help to refer to Figure 22 to keep track of the process rankings.) In that case, we would have found it impossible to fit the observed data without drastically revising the hierarchy of vowel loss processes. And, since we had more data to support that hierarchy than for cluster simplification, it would have been reasonable to reject the hypothesized rank of #CCV> #CV. (However, if #CCV > #CV had been ordered higher, a proportionate increase in V → ϕ/#C would not have changed the vowel loss ordering, since loss in this context was assumed to be the most likely.)

The probability of occurrence of the states with no initial vowels appears also to be of the correct order of magnitude (0.002). If it were much larger, instances of languages of that type should be known. If it were much smaller, languages such as OYKANGAND would be unlikely to approach the type so closely. Thus the order of magnitude difference between the processes V → ϕ/#__C (0.06) and #CV > #V (0.0004) is approximately correct.

The model predicts probabilities of the order of 10^{-6} or less for languages without medial vowels. This value is determined mainly by the process VCV > VV (0.002) and the various processes of epenthesis. There is no reason that the state probabilities could not be even less, for there is no lower bound to likelihood of VCV > VV. Under our

assumptions, however, it is unlikely that the state probabilities are greater than about 10^{-4}, since the likelihood of VCV > VV is bounded above by #CV > #V and the epenthetic processes occur sufficiently often that the hypothesized likelihoods are probably rot overstated by more than a factor of 100.

The addition of the word combination processes has improved the model's representation of observed asymmetries of distribution. Notably, the relative values of the states with no medial clusters ($\bar{M} = 1$) are much improved. A shift toward final clusters for languages with 3-segment medial clusters ($\bar{M} = 3$) may also be observed. The importance of word combination in shaping syllable structure is therefore supported.

It should be evident, however, that the model is still not a satisfactory fit to the observed distribution, even with reasonable allowance for distortions of sampling. For $\bar{I} = 1$, the values for $\bar{M} = 1$ are too low, and for $\bar{I} = 2$, they are too high. The predominance of medial clustering for languages with initial and final clusters is inadequately expressed. In other words, while addition of the word combination processes has produced asymmetries of the correct general nature, its effect was incorrect in detail or quantitatively insufficient, or both.

More juggling of the parameters, notably a still larger increase in the effect of word formation, would improve the picture somewhat. I did not feel that this step was worthwhile. For one thing, the value WF = 4 means that the proportional effect of word formation on medial cluster formation is 5 or 6 times that of the phonetic processes. A much greater proportion may be realistic. Yet without some independent evidence, I am unwilling to push this proportion much higher. Given our initial assumptions, the model has been carried as far as is prudent. Further progress requires re-examination of the assumptions and independent verification where possible.

3.2.4 Future directions. The original impetus behind the Markov chain model was to find a nondeterministic explanation for the rarity or nonoccurrence of certain states. The model has turned out to have much greater potential. Some secondary benefits have already accrued. Through the model, the processes that change syllable structure are related, so that hypotheses about one can to some extent be confirmed by evidence about others. As the model is refined, the relationship will become closer. Another benefit, as we shall see shortly, is that the effort to improve the model suggests interesting hypotheses about linguistic processes that deserve study.

In refining the model of clustering, a prudent first step would be backwards, to shore up the foundations. The assumptions behind the values for cluster simplification and word combination need firmer support. This could be provided by independent studies, both by means of case studies of the development of syllable structure over long periods of time and by applying probabilistic state-process models to the study of subtypes.

Although such work would doubtless lead to the revelation and correction of errors in my assignment of relative likelihoods to the processes, I have little hope that the major deficiencies of the model could thereby be remedied. If this is so, it follows that a basic assumption about the processes of syllable structure made earlier (Section 3.3.3.), namely that the relative likelihoods of processes are independent of the states, should be rejected. (Indeed, we will be led to question below even the premise that the processes are independent of the past history of the languages they act upon.)

So we are faced with the prospect of determining how processes depend upon the states and by what principles they are governed. It looks as though the dependencies are something like this. The likelihood of initial clustering is greater for languages with final consonants

than it is for languages without final consonants, and greater still for languages with final clusters. The likelihood of final clustering is greater for languages with initial clusters than it is for languages without.[25]

These relations are not proposed seriously as a substitute hypothesis for independence. They are much too strong, in the sense that even though they doubtless could explain the present data, this would tell us nothing about the mechanism involved, which is what is really important. The notion that the form of one end of a word directly affects what happens at the other end is not very plausible. There must be an intermediate linking mechanism.

Some candidates came to mind. What about the role of word accentual systems? Loss of consonants and vowels is closely connected with accentual position. In some cases, too, it appears to be capable of producing an initial-final dependency. Unfortunately, the dependency is the inverse of the one postulated. For instance, given the formation of initial clusters in a language, we may presume the chance of the language possessing initial stress is less than for languages in general. Hence, it is more likely than usual to have a final stress system, which would diminish the chance of forming final clusters. The exact way in which accentual systems interact with syllable structure processes is complex, and it is not possible to judge offhand the magnitude of its effect. In any case, it can hardly be the linking mechanism needed.

A more promising direction may be to explore the possible tendency of affixes to be especially vulnerable to reduction in substance. Assuming affixes to be derived ultimately from combination of independent

---

[25] 'Initial clustering' and 'final clustering' stand here for the combined effect of processes of formation and simplification. The stated dependencies could hold for one or more of them.

words, this would imply that medial clusters formed by word combination would be more likely to emerge as initial or final clusters than medial clusters formed by phonetic processes. There might even be a preference, once such elements are reduced to single segments, for them to resist loss. Either preference would result in dependencies of the kind postulated, but only further work could tell whether they are of sufficient magnitude to be significant.

One more possibility should be mentioned. There may be a distinction between languages which are 'active' with respect to syllable structure and those which are passive. Languages such as ARABIC, KORYAK, and YOKUTS are called active because their syllabic canons would be violated without a set of sandhi rules which maintain the language's structure. On the other hand, a language like ENGLISH is passive because combinations of forms do not violate its canons and and any sandhi rules do not bear directly on its syllable structure. It may be worth investigating whether there is a difference in the diachronic behavior of such languages and to what extent they exhibit preferences for certain syllable types.

These last two possible mechanisms both involve word combination, if the term is extended to include morphological structure. This reinforces somewhat the earlier conclusion that word combination was a major, even dominant, process in the dynamics of syllable structure.

There may be some readers who would object to the Markovian assumption that processes are independent of a language's history. They might claim, for example, that a language which has recently developed s + stop initial clusters is much less likely to lose them and return to its original state than a language possessing at some previous point many initial clusters of which it has lost all but those in s + stop. This may be so, but it will be difficult to establish. One approach would be to establish state-process models within subtypes. It is not enough

to point to examples of the repeated occurrence of processes in a language over a long period. Those who are convinced by such evidence should read Feller's (1968, Chapter III) lucid account of how Markovian processes can give rise to periods away from equilibrium of seemingly improbable duration. The question is of considerable interest. As for the validity of the overall model, unless the dependencies are great, the assumption of the Markov property will still produce a satisfactory approximation. This problem will also arise if the above-mentioned preferences for loss in affixes and retention of their reduced form are incorporated into the model, since then some process likelihoods will depend upon whether medial clusters are phonetic or combinatorial in origin, and therefore will naturally depend upon the distribution of the two types in the states of origin.

3.3 A Markov state-process model of vowel onset. The formation of syllables with vowel onsets can be investigated by means of a process-state model. The vowel onsets may occur initially or medially. Medially, they usually occur when two syllable nuclei abut in hiatus. This is the case considered here. The other possibility arises when syllable division of the type -C.V- occurs. It was discussed in Section 2.4.1.

The language types, therefore, that concern us are given by the values of the two parameters I and M, corresponding to a crossclassification of languages with and without hiatus and with and without initial vowels. The distribution of these types was given above in Figure 4. Here I use the distribution obtained by applying the corrections for oversampling discussed in Section 3.2.3.

I = 1 (no initial vowels)	I = 0 (initial vowels)	
25	31	M = 1 (no hiatus)
17	61	M = 0 (hiatus)

Figure 24. Corrected number of languages occurring according to vowel onset types.

The processes that link the states are shown in Figure 25. The usual symbolism is used, the digits identifying the states standing for for values of $\underline{I}$ and $\underline{M}$ respectively. State 10 contains languages with no initial vowels ($\underline{I} = 1$) and with hiatus ($\underline{M} = 0$), for example. The asterisked processes are those that must be carried to completion to effect a change of state. A language must lose all its initial vowels to move from $\underline{I} = 0$ to $\underline{I} = 1$. The only process of word combination that is relevant is the combination of two words joined by a V#V sequence, forming a sequence of two nuclei. It naturally occurs only with languages possessing initial vowels. The general nature of the processes was discussed in Section 3.1.

It should be apparent that the strategy of hypothesizing relative likelihoods of processes is not useful here. There is simply no evidence for most of these likelihoods of anything like the weight of the observed distribution of languages. The approach to use here is to work backward from the state proportions and see what inference can be made about the relative likelihoods of the processes.

The representation of processes and states in Figure 26 will be used. The proportion of languages occurring in each state is given by X_1, X_2, X_3, and X_4. The relative likelihoods of the phonetic processes between each state are p_0, p_1, q_0, and q_1; w is the relative likelihood of hiatus occurring by word combination.

The equations relating the state probabilities and transition probabilities are

$$
\begin{aligned}
-(p_1+q_1)X_1 + \quad p_0X_2 + \quad q_0X_3 \quad\quad\quad &= 0 \\
p_1X_1 - (p_0+q_1+w)X_2 \quad + \quad q_0X_4 &= 0 \\
q_1X_1 \quad - (p_1+q_0)X_3 \quad + \quad q_0X_4 &= 0 \\
(q_1+w)X_2 \quad + \quad - (p_0+q_0)X_4 &= 0
\end{aligned}
\tag{64}
$$

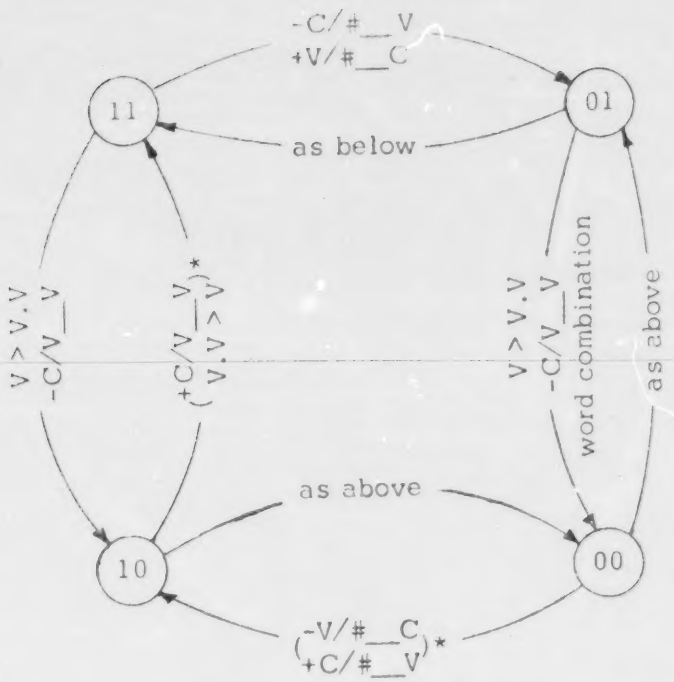


Figure 25. Processes affecting word-initial vowels and hiatus.

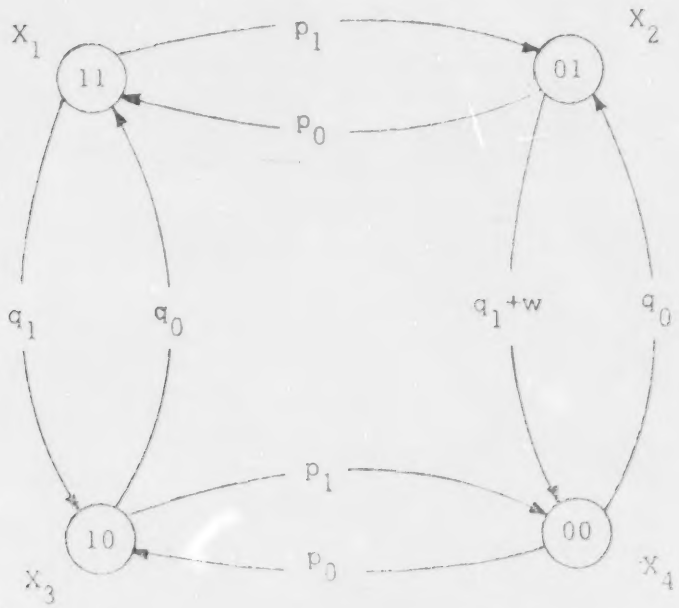


Figure 26. Transition and state probabilities for the Markov chain model.

The equations can be rearranged with the transition probabilities as unknowns.

$$
\begin{aligned}
+X_2p_0 - X_1p_1 + X_3q_0 - X_1q_1 &= 0 \\
-X_2p_0 + X_1p_1 + X_4q_0 - X_2q_1 - X_2w &= 0 \\
+X_4p_0 - X_3p_1 - X_3q_0 + X_1q_1 &= 0 \\
-X_4p_0 + X_3p_1 - X_4q_0 + X_2q_1 + X_2w &= 0
\end{aligned} \tag{15}
$$

General constraints on the transitions can now be obtained. Adding the first and third equations, we have

$$\frac{p_1}{p_0} = \frac{X_2 + X_4}{X_1 + X_3} \tag{16}$$

Adding the first and second (or third and fourth) equations gives

$$(X_3 + X_4)q_0 - (X_1 + X_2)q_1 - X_2w = 0, \quad \text{or} \tag{17}$$

$$\frac{q_1}{q_0} + \frac{X_2}{X_1+X_2}\ \frac{w}{q_0} = \frac{X_3 + X_4}{X_1 + X_2} \tag{18}$$

Note that the process of word combination is the only source of asymmetry here. If $w = 0$, an equation for q_1/q_0 similar to that for p_1/p_0 is obtained, and the parameters I and M will be uncorrelated. This follows from the assumption that the processes concerning initial position are not affected by the presence or absence of hiatus, and vice-versa.

The final[26] constraint provided by the model links the hiatus transitions and the initial vowel transitions:

$$\frac{q_1}{q_0} = \frac{X_3}{X_1} - \frac{X_1X_4 - X_2X_3}{X_1(X_1 + X_3)}\ \frac{p_0}{q_0} \tag{19}$$

---

[26] No more than three independent constraints can be obtained from the set of four equations (15) because the sum of the first two equations less the sum of the last two equations yields the tautology $0 = 0$.

The observed values for X_1, X_2, X_3 and X_4 are now substituted in the equations (16), (18), and (19).

$$(20) \quad \frac{p_1}{p_0} = 2.1$$

$$(21) \quad \frac{q_1}{q_0} + .54 \frac{w}{q_0} = 1.5$$

$$(22) \quad \frac{q_1}{q_0} = .72 - .93 \frac{p_0}{q_0}$$

I can now draw a number of conslusions about the processes. From (20) I conclude that

(23) The loss of all a language's initial vowels (or prothesis of a consonant before them) is significantly less likely than a language with no initial vowels acquiring them by consonant loss or vowel prothesis.[27]

The processes affecting hiatus are more complicated. In Figure 27 the conditions (21) and (22) are graphed, as well as the ratios w/q_1 and p_1/q_1. The latter are derived from (20), (21), and (22). Since none of the ratios may be negative we see that q_1/q_0 must be less than 0.72. Otherwise, a wide range of parameter values can satisfy the data.

Let us first consider the ratio p_1/q_1. It represents the relative strength of the processes that introduce initial vowels and hiatus, mainly loss of initial consonants and loss of medial consonants. I suspect that most linguists familiar with phonetic change would say that medial intervocalic loss is more frequent than loss before initial vowels.

---

[27] It would be gratifying if this confirmed the assumptions used for the cluster model above. A comparison is not possible, however. There we were concerned with any instance of vowel loss, and the total loss of initial consonants, just the opposite of the present case.

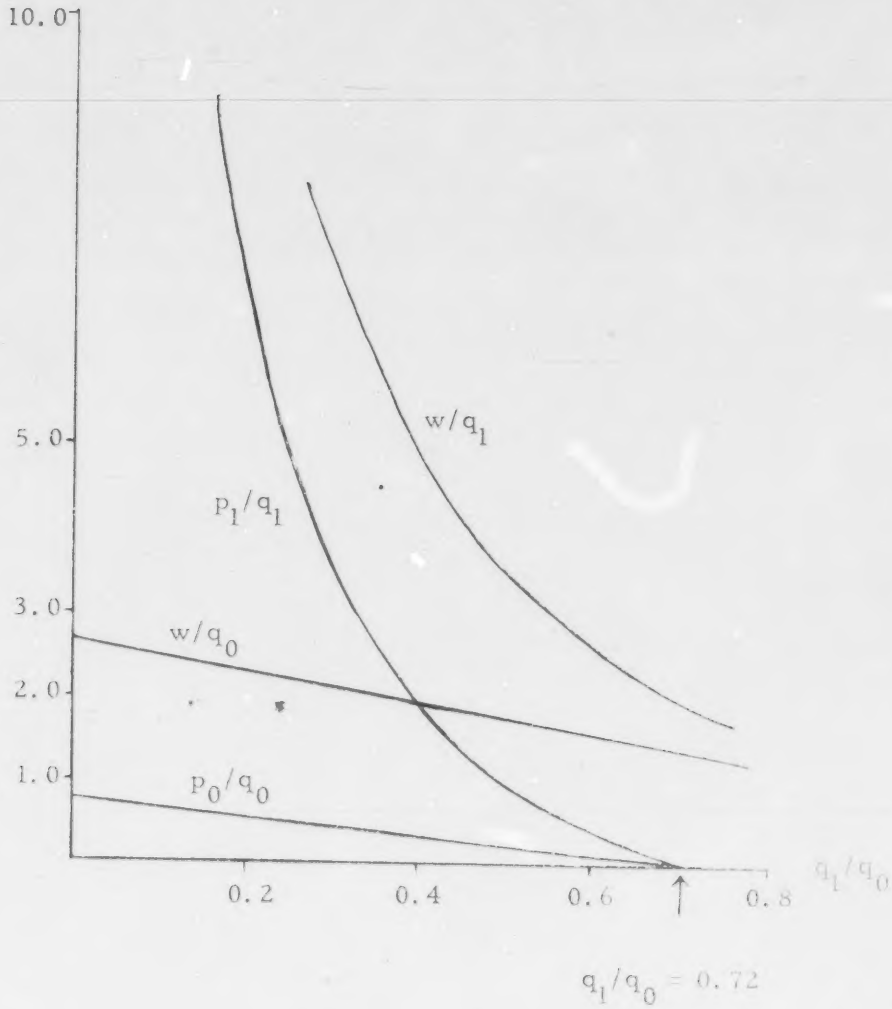


Figure 27. Constraints on the parameters of the vowel onset model.

But we do not have to depend entirely upon intuitions. Among the language descriptions of the data, there are frequent references to recent changes, or to alternations that point to changes in progress. Without much more elaborate precautions, this sort of data is probably as good a sample of relative frequency of occurrence of phonetic change as we can expect.

The data show medial loss to be much more frequent: 23 instances of medial consonant loss between vowels were found, to 8 instances of initial vowel loss. I rush to caution the reader that the extent of the preference is probably not applicable to the immediate situation. Loss of an intervocalic consonant does not necessarily lead to vowel hiatus. A long vowel or diphthong may be formed immediately by readjustment of the syllable structure. Such cases could be treated conceptually as creation of hiatus and its immediately subsequent destruction, although this does not attract me particularly. Among the 23 cases of consonant loss, 10 resulted in hiatus, 7 did not, and the outcome of 6 was undeterminable. It appears that the following conjecture is warranted.[28]

(24) The loss of intervocalic consonants is a more frequently occurring process than loss of initial consonants; hiatus-producing intervocalic loss is less frequent but probably still more frequent than initial loss.

Accepting this conjecture as part of our tissue of hypotheses, the parameters of the process are much more greatly restrained. It implies

---

[28] Although there is little purpose in sniping endlessly at the data, since it is all that we have, I mention another defect in the application of this data. The processes of the model are changes that take place in languages that do not possess hiatus or initial vowels; the data on the frequency of occurrence of the processes is not so restricted. Implicit in my use of the data is the unverified assumption that the relative frequencies of occurrence in the two situations are not greatly different.

that $p_1/q_1 < 1$, so that the process must have relative transition values found on the right of the graph in Figure 27. If we took the exact values seriously (which we must not do; since the numbers are not characteristics of some physical process, but estimates whose bias and variance are unknown, we must be cautious) q_1/q_0 must be greater than 0.5. I feel the following conclusions about q_1/q_0 are justified.

(25) The phonetic processes of creation and destruction of hiatus, mainly loss of intervocalic consonants opposed to intervocalic epenthesis ($\phi \rightarrow C/V__V$) and nuclear fusion (V.V > V: elision, contraction, diphthongization), have likelihoods of occurrence that are of the same magnitude, although intervocalic consonantal loss is probably the less frequent of the two.

The model indicates p_0/q_0 is less than 1.0 in all circumstances, and less than 0.3 if $p_1 < q_1$. I conclude that

(26) The complete elimination of a language's initial vowels, whether by loss or by consonant prothesis, is less likely than the complete elimination of a language's sequences of syllable nuclei.

In languages with initial vowels hiatus may arise either by phonetic change or by word combination. The ratio w/q_1 indicates the relative weight of the two processes. The word combination component is dominant, having roughtly 2 to 4 times the likelihood of the phonetic processes.[29] That a similar conclusion was reached for the clustering

---

[29] The results could easily be extended to cover vowel-initial syllables that occur in -C.V- division. Since -C.V- division does not usually occur without a constituent boundary, it is reasonable to suppose that most such cases reflect an original word combination. The same model could be applied to data that included occurrence of C.V division, and the increased value of the word combination likelihood would be an indication of the strength of the process VC#V > C.V (as distinct from the process with syllabic reorganization -- VC#V > V.CV).

model is interesting but probably not particularly significant, since the phonetic processes are very different in the two cases.

Using a Markov chain to model vowel onset dynamics has been fruitful. First, it has been demonstrated that a model with state-independent phonetic processes plus one additional transition representing word combination is rich enough to explain the observed dependence of hiatus and presence of initial vowels. It is also the simplest possible model without state-dependent processes.

Second, the model revealed relationships among the process likelihoods and state probabilities that could be exploited to predict rankings and magnitudes of the process likelihoods. Future work may verify or discredit some of the assumptions or conclusions, but no longer need they be treated as isolated problems.

I expect that the model itself will be improved. There are already a couple of hints of inadequacy. I stated earlier that the distribution of languages appeared to be largely independent of the clustering parameters. One apparent dependency, for which I have no explanation, did turn up. Languages without initial vowels occurred in the sample less often with hiatus if they had initial clusters than if they did not.[30]

On the other hand, a dependency that can be predicted from the models does not appear. Recall that word combination is the only

---

[30] For $\bar{I} = 1$ the uncorrected distribution of languages was $X_1(1) = 12$, $X_2(1) = 18$, $X_3(1) = 16$, $X_4(1) = 33$; for $\bar{I} = 2$ it was $X_1(2) = 16$, $X_2 = 14$, $X_3(2) = 2$, $X_4(2) = 33$.

$\bar{I} = 1$		$\bar{I} = 2$		Total	
12	18	16	14	28	32
16	33	2	33	18	66

The left-hand columns in each distribution are languages without initial vowels; the top rows are languages without hiatus.

source of the correlation between I and M. Now, if the model used in Section 3.2.2.5. is correct, then the likelihood of forming V#V sequences should be greater for languages that have no final consonants. The excess of initial-vowel languages with hiatus should be greater for languages with no final consonants. The data manifestly do not reflect such a tendency.[31]

4. Conclusions

One accomplishment of the investigation is that we are able to modify and extend the generalizations of Jakobson and Greenberg. In general, our survey of syllable types confirmed the findings of Greenberg. All languages in the sample possessed CV syllables and at least one other type. However, the implication that possession of VC syllables implies the possession of V syllables does not necessarily hold. In terms of the subset typology, known occurring syllable structures of the basic syllable types are the following: CV, V—ELLICE; CV, CVC – TIBETAN; CV, CVC, V – ARABELA; CV, CVC, VC – NORTHERN PALESTINIAN ARABIC; and CV, V, CVC, VC – BURMESE.

The type with inventory CV, V, VC is missing, although its occurrence is predicted by Greenberg's generalizations.[32] It appears that

---

[31] The (uncorrected) distribution of languages for the values $\bar{F} = 0$, $\bar{F} = 1$, and $\bar{F} = 2$ is given below:

$\bar{F} = 0$		$\bar{F} = 1$		$\bar{F} = 2$	
5	8	17	15	6	9
4	16	10	31	4	19

The measure of correlation $(X_1X_4 - X_2X_3)/N^2$, where N is the number of languages in each case, was .044 for $\bar{F} = 0$, .071 for $\bar{F} = 1$, and .014 for $\bar{F} = 2$.

[32] Such a language could only occur by the unlikely circumstance that the marked syllabification C.V only occurred in a sequence .VC.V, or final consonants occurred only in the sequence V.VC#.

in the terms of the subset typology, the generalizations of Greenberg should be modified to read as follows (my additions to the original statement are underlined).

(27) '...if syllables containing sequences of n consonants in a language are to be found as syllabic types, the sequences of n-1 consonants are also to be found in the corresponding position (prevocalic or postvocalic) except that CV → V ...' and VC → V'... do not hold' and V(C) → CV(C). (Greenberg, Osgood, and Jenkins 1963, xxv)

From the point of view of the parametric typology, in which syllable structure is defined over a larger domain, a number of basic generalizations were made. They are repeated here for convenience.

(1) All languages have initial margins.
(2) All languages (except monosyllabic ones) have medial margins.
(3) All languages have medial margins of one segment.
(4) All languages have some open final syllables.
(5) If a language has initial or final clusters, then it also has medial clusters.
(6) Languages with uniform syllable division do not have VC.V division.

What is the significance of these generalizations? To the extent that they can be explained by regularities of diachronic process, I have concluded that they are not the result of deterministic constraints or processes. The statements (27) and (1) through (6) are stated in terms of prohibition of occurrence, but they signify in fact the rarity, moderate to extreme, of certain kinds of syllable structures. That this view is correct is substantiated by the existence of languages that very nearly exemplify some of the prohibited types. Further, it appears that the implication (5) is but an extreme part of a general correlation between medial and terminal margin lengths.

Let us see how the universality of the CV syllable can be explained in these terms. First, though, its synchronic predominance should be placed in perspective. The predominance does not appear to be based upon the inherent nature of the syllable itself. After all, CV syllables are probably more likely to disappear by vowel loss than CVC syllables. Such a preference is represented in the phonology of YOKUTS, which includes a rule $\overset{-\text{long}}{\text{V}} \rightarrow \phi / \text{VC}__\text{CV}$ (Kisseberth 1970). If the type CV excludes $\text{C}\bar{\text{V}}$ and $\text{CV}\hat{\text{V}}$, we find languages where its occurrence is extremely limited -- WIYOT, only word finally (Teeter 19 7); NAVAHO, only before syllabic n or a final long vowel (Hoijer and Sapir 1967). Nor are CV syllables always most frequent, even when their privilege of occurrence is not greatly constrained; in a sample of 505 KOLAMI syllables CVC syllables were more frequent, and other examples could probably be found.

Why then do all languages have CV syllables? (Or what makes them so rare that none are known?) The answer, I think, is that all such languages are violations of one or more of (6) and (1) through (4); their occurrence is disfavored by the strength of the processes leading away from the type compared to those leading to it.[33] A language with just VC syllables violates (1), (4), and (6); with just CVC, (3) and (4); etc. The most likely kind of syllable structure without CV syllables would involve some combination of CVC syllable types with other types. Such a language would have had to have lost all its final vowels (unless the final syllable was V). In medial position, only the sequences -C.C-, -C.V-, or -.V.V- could occur. A rough estimate of the likelihood of occurrence of such languages could be made using the techniques of the previous sections, but it is evidently very small.

---

[33] A processual explanation for generalization (6) concerning syllable division is not presumed.

What about monosyllabic languages? The medial constraints do not hold in their case. Do we therefore expect a CVC, VC language to be more likely in this case? I do not know. The entire process of development of word-length constraints may be involved.

Apart from its substantive results, the investigation has extended the scope of dynamic comparison as a typological technique (Greenberg 1966a, 1969). The discussion of the CV syllable illustrates the path that the investigation has taken. A synchronic generalization about occurrence and nonoccurrence of types was considered. In terms of occurrence and nonoccurrence, the generalizations did not appear to be lawlike, for they had no processual consequences. Yet when they were treated as statements of rarity of occurrence, it was possible to relate them to the relative likelihoods of processes. Furthermore, the same probabilistic techniques opened the way to the explanation of the relative proportions of languages found among commonly occurring types.[34]

Although the present use of probabilistic models was directed to the explanation of synchronic observations, the technique's greatest promise probably does not lie there. The opportunity it provides to explore the relative strengths of diachronic processes through its power to interrelate observations about states and process is an exciting prospect. It also appears to afford a means of attack upon some fundamental questions of linguistic change grouped under the headings of 'drift', 'teleological nature', etc. Extended application of the technique, in whatever direction, will require much more than just the construction of models. The only contribution of the model, although

---

[34] Synchronic regularities of relative proportion can also sometimes be explained by deterministic diachronic cases; an example is the hypothesis (1-12) regarding the origins of syllabic nasals. (Bell (1970a).

it is central, is to relate otherwise isolated evidence. The evidence must still be obtained. For example, the data of this investigation, being of shallow time depth, did not do justice to the power of the models constructed. In particular, the data did not bear witness to the nature of word combination and the nature of change of state by process culmination (e.g. loss of all final consonants). Case studies of change over long periods are also needed, such as those of Martinet (1964) and Shevelov and Chew (1969). A better knowledge of the nature of language samples will also be desirable, and may possibly be attained by studies of areal influence and studies of the evolution of new languages and the extinction of old.

APPENDIX

THE DATA

Introduction. The principal source of data for the investigation is a basic sample of 144 languages. The languages were chosen according to three main criteria. The availability of an adequate description was naturally necessary. Secondly, the languages were chosen to obtain coverage of the major genetic families and language areas. Some effort was made to keep the number of languages sampled roughly in balance with respect to families, but this goal was violated in several ways. Some languages were sampled because of availability of descriptions, e.g. a volume containing descriptions from a given area. In several instances, the sample included quite closely related languages or dialects which were added for the purpose of obtaining diachronic perspective. A third criterion was the exemplification of a rare or interesting type. The inclusion of OYKANGAND, an Australian language of unusual syllable structure, is an example.

Partial descriptions of a further 25 languages were used to illustrate particular phenomena.

Comparative and historical studies were a secondary source of data. Examples are the CHINESE dialect dictionary Hanyu Fangyin Zihui, and Dobson's English Pronunciation 1500-1700.

A.1 The languages. Following are listed alphabetically the languages that were used in the investigation. The asterisked languages do not belong to the basic sample, but were used with reference to a particular phenomenon. Sources consulted for the language are given in the right-hand column.

Languages Investigated	Sources Consulted
Acoma	Miller (1965)
Aguacatec	McArthur & McArthur (1956)
Akan	Schachter & Fromkin (1968), Welmers (1946)
Alabaman	Rand (1968)
Amharic	Cohen (1936)
Amuzgo	Bauernschmidt (1965)
Apinayé	Burgess & Ham (1968)
Arabela	Rich (1963)
*Eastern Arabian Arabic	Johnstone (1967a, 1967b)
Arapaho	Salzmann (1956)
Auca	Saint & Pike (1962), Pike (1964)
Balese	Vorbichler (1965)
Baure	Baptista & Wallin (1968)
Bashkir	Poppe (1964)
*Basque	Grammont (1933)
Bengali	Ferguson & Chowdhury (1960)
Berti	Petraček (1965)
Bilin	Palmer (1957, 1958, 1960)
Burmese	Burling (1967), Cornyn (1945) Stewart (1955), Zlatoverxova (1966)
Bushong	Vansina (1959)
Campa	Dirks (1953)
*Cantonese	Cheng (1968)
Carib	Hoff (1968)
Cashinahua	Kensinger (1963)
Cayapa	Lindskoog & Brend (1962), Moore (1962)
Colorado	Moore (1962)

Languages Investigated	Sources Consulted
Chácobo	Prost (1967)
Cham	Blood (1967)
Chipaya	Olsen (1967)
Chukchee	Bogoras (1922), Skorik (1961)
Cofan	Borman (1962)
*Cora	McMahon (1967)
Czech	Komárek (1958), Kučera (1961)
Dan	Bearth & Zemp (1967)
*Danish	Lauritsen (1968)
Dogon (Kamba)	Kervran & Prost (1969)
Dogon (Toro)	Calame-Griaule (1968)
*Dutch	Shetter (1958)
Egyptian Arabic	Harrell (1956), Harrell et al. (1963), Lehn (1963), Mitchell (1956)
Eskimo (Inupik)	Holtved (1952), Swadesh (1946), Swadesh (1951, 1952a, 1952b, 1952c), Ulving (1953)
Eskimo (Yupik)	Swadesh (1951, 1952a, 1952b, 1952c)
Ellice	Milner (1958)
English	Bailey (1967, 1968c), Bailey & Milner (1967), Dobson (1968)
Esako	Laver (1967)
Estonian	Harms (1962)
Fijian	Scott (1948)
*French	Fouché (1959), Pulgram (1960-61, 1965a)
*French (Gascon)	Grammont (1933)
Fulani	Stennes (1967)
Galla	Andrzeyewski (1957), Moreno (1939)
Garo	Burling (1961)

Languages Investigated	Sources Consulted
Gbeya	Samarin (1966)
Georgian	Tschenkéli (1958)
*German	Mohr (1970)
*Grebo	Innes (1966)
*Greek (Classical)	W.S. Allen (1968), Jakobson (1962), Ward (1946)
*Greek (Modern)	Ward (1946)
Guarani	Gregores & Suárez (1967)
Gugu-Yalanji	Oates (1964), Oates & Oates (1964)
Gujarati	Cardona (1965), Pandit (1957, 1958)
Halang	Cooper & Cooper (1966)
Hakka	Henne (1965)
Hausa	Hodge (1947), Greenberg (1941)
Hottentot (Nama)	Beach (1938)
Huichol	Grimes (1959), McIntosh (1945)
Hungarian	Tompa (1968)
Ifugao	Newell (1956)
Ignaciano	Ott & Ott (1967)
Ijo	Williamson (1965)
*Ipuriná	Noble (1965)
Irish (Ring County)	Breatnach (1947)
Itonama	Liccardi & Grimes (1968)
Japanese (Standard)	Bloch (1950), Martin (1952, 1959), McCawley (1965, 1968), Wenck (1954)
Jeh	Cohen (1966), Gradin (1966)
Kaingáng	Henry (1935), Wiesemann (1964)
Kannada	Hiremath (1961), Nayak (1967)
Karačay	Hebert (1962)
Kashmiri	Kachru (1969), Kelkar & Trisal (1964), Morgenstierne (1941), Varma (1964)

Languages Investigated	Sources Consulted
Kayan	Cubit (1964)
Khasi	Henderson (1965), Pinnow (1959), Rabel (1961)
Khmer	Henderson (1952), Jacob (1960, 1966), Noss (1966)
Khmu	Smalley (1961)
Kikuyu	Barlow (1960)
Kolami	Emeneau (1955)
Korean	Martin (1951)
Koryak	Kreinovič (1958b)
Koya	Tyler (1969)
Kurukh	Pinnow (1964-65)
*Lahu	Burling (1967), Matisoff (1969)
*Lalana Chinantec	Rensch (1966), Rensch & Rensch (1966)
*Latin	Allen (1967)
Latvian	Abele (1924), Matthews (1959), Zeps (1962)
Lendu	Tucker (1940)
Luganda	Cole (1967), Meeussen (1955), Tucker (1962)
Luo	Gregerson (1961)
*Machiguenga	Noble (1965)
Madurese	Stevens (1968)
Maidu	Shipley (1956), Uldall (1954)
Mandarin Chinese	Chao (1948, 1968), Hockett (1947)
Margi	Hoffman (1963), Ladefoged (1968)
*Maru	Burling (1966, 1967), Lyovin (1968)
Mazahua	Pike (1951), Spotts (1953, 1956)
Mazateco	Pike & Pike (1947)

Languages Investigated	Sources Consulted
Maori	Biggs (1961), Haslev (1967), Hohepa (1967), Holmer (1966), Krupa (1966)
Middle Wahgi	Luzbetak (1956)
Minnan Chinese (Gaoxiong)	Tung (1968)
Totontepec Mixe	Crawford (1963)
Mongolian	Poppe (1970)
Navaho	Sapir & Hoijer (1967)
Njungar	Douglas (1968)
Norwegian (Standard)	Næs (1965), Popperwell (1963)
*Norwegian (Trondheim)	Sivertsen (1968), Vanvik (1966)
Nukuoro	Carroll (n.d.)
Osage	Wolff (1952)
Oykangand	Dixon (1970), Sommer (1969, 1970)
Pacoh	Watson (1964)
Palaung	Shorto (1960)
North Palestinian Arabic	Blanc (1953)
Pame	Gibson (1956)
Pashto	MacKenzie (1958), Penzl (1955)
Paresí	Rowan (1967)
*Persian	Grammont (1933)
Piro	Matteson (1965), Matteson & Pike (1958)
Eastern Popoluca	Kalstrom & Pike (1968)
Quechua	Orr (1962a, 1962b), Reyburn (1954)
Quiotepec Chinantec	Robbins (1961)
Russian	Avanesov (1956, 1968), Isačenko (1947), Kučera & Monroe (1968) Shapiro (1968)
Sangir	Maryott (1961)
*Sanskrit	Allen (1953)

Languages Investigated	Sources Consulted
Sara (Majingay)	Bouquiaux (1963), Hallaire & Robinne (1955-59)
Secoya	Johnson & Peeke (1962)
Sedang	Smith (1968)
Sentani	Cowan (1965a)
Shilha	Applegate (1958)
Sierra Nahuat	Key & Key (1953)
Sinhalese	Coates & de Silva (1960), de Silva (1959)
Siona	Wheeler & Wheeler (1962)
Songhai	Prost (1956)
Southern Paiute	Harms (1966), Sapir (1930)
Spanish (Standard)	Navarro Tomás (1957)
Squamish	Kuipers (1967)
Susu	Houis (1963), Sangster & Faber (1968)
Swahili (Bravanese)	Whiteley (1965)
Swahili (Standard)	Greenberg (1947), Polomé (1967) Tucker & Ashton (1942), Tucker (1946-47)
Tagalog	Bowen (1965), Lopez (1965)
Tamazight	Abdel-Massih (1968)
Telefol	Healey (1964)
Tetelcingo Nahuatl	Pittman (1961)
Thai	Abramson (1962), Fowler & Isarena (1952), Noss (1964)
Tibetan	Miller (1935)
Totonaco	Aschmann (1946)
Trique	Hamp (1954), Longacre (1952, 1955, 1959)
Tswana	Cole (1955)

Languages Investigated	Sources Consulted
*Turkish (Bulgarian)	Eckmann (1962), Kakuk (1958)
Turkish (Standard)	Swift (1963)
Usila Chinantec	Skinner (1962)
Walapai	Redden (1966)
Warao	Osborne (1966)
Western Desert (Njangumada)	O'Grady (1957)
Western Desert (Ooldea)	Douglas (1954-55, 1964)
Wik-Munkan	Sayers (1974)
Wiyot	Teeter (1964)
*Yahad	Cowan (1965)
Yakur	Bendor-Samuel (1969)
*Yay	Gedney (1965)
*Yokuts	Kisseberth (1970)
Yoruba	Bamgbose (1966)
Yukaghir	Kreinovič (1958a)
*Isthmus Zapotec	Pickett (1951, 1960)
Zoque	Wonderley (1951)

A.2. Genetic distribution of the languages. The languages in the basic sample are listed below according to their atfiliation with the world's major language families. Three families in the classification are not widely accepted. The classification of Indo-Pacific languages is discussed in Greenberg (to appear). The Eurasiatic and Amerind families have been proposed informally by J.H. Greenberg. Eurasiatic is composed of the branches Uralic; Altaic; Gilyak; Yukaghir; Chukchee, Kamchadal, Koryak; Ainu, Japanese, Korean; and Aleut and Eskimo. All the indigenous language families of America except Na-Dene belong to a common Amerind stock, although many details of sub-classification within Amerind are still undecided.

Major Family	Number in Sample	Languages
Indo-European	(12)	Bengali, Czech, English, Gujarati, Irish, Kashmiri, Latvian, Norwegian, Pashto, Russian, Sinhalese, Spanish
Caucasian	(1)	Georgian
Eurasiatic	(13)	
Uralic		Estonian, Hungarian
Altaic		Bashkir, Karačay, Mongolian, Turkish
Other branches		Chukchee, Eskimo (Inupik), Eskimo (Yupik), Japanese, Korean, Koryak, Yukaghir
Dravidian	(4)	Kannada, Kolami, Koya, Kurukh
Austroasiatic	(8)	
Mon-Khmer		Halang, Jeh, Khmer, Pacoh, Sedang
Other branches		Khasi, Khmu, Palaung
Austronesian	(11)	Cham, Ellice, Fijian, Ifugao, Kayan, Madurese, Maori, Nukuoro, Sangir, Tagalog, Thai
Indo-Pacific	(3)	Sentani, Telefol, Middle Wahgi
Australian	(6)	Gugu-Yalanji, Njungar, Oykangand, Western Desert (Njangumada), Western Desert (Ooldea), Wik-Munkan
Sino-Tibetan	(6)	
Sinitic		Hakka, Mandarin, Munnan Chinese (Gaoxiong)
Tibeto-Burman		Burmese, Garo, Tibetan
Na-Dene		Navaho

Major Family	Number in Sample	Languages
Amerind	(40)	
Penutian		Aguacatec, Chipaya, Maidu Totontepec Mixe, Totonaco, Zoque
Aztec-Tanoan		Huichol, Sierra Nahuat, Southern Paiute, Tetelcingo Nahuatl
Otomangean		Amuzgo, Mazahua, Mazateco, Pame, Eastern Popolocan, Quiotepec Chinantec, Trique, Usila Chinantec
Macro-Chibchan		Cayapa, Colorado, Itonama, Warao
Ge-Pano-Carib		Apinayé, Carib, Cashinahua, Chácobo, Kaingáng
Andean-Equatorial		Arabela, Auca, Baure, Campa, Cofan, Guarani, Ignaciano, Paresí, Piro, Quechua, Secoya, Siona
Other branches		Acoma, Alabaman, Arapaho, Osage, Squamish, Walapai, Wiyot
Afro-Asiatic	(9)	Amharic, Bilin, Egyptian Arabic, Galla, Hausa, Margi, North Palestinian Arabic, Shilha, Tamazight
Congo-Kordofanian	(17)	Akan, Bushong, Dan, Dogon (Kamba) Dogon (Toro), Esako, Fulani, Gbeya, Ijo, Kikuyu, Luganda, Susu, Swahili (Bravanese), Swahili (Standard), Tswana, Yakur, Yoruba
Nilo-Saharan	(6)	Balese, Berti, Lendu, Luo, Sara (Majingay), Songhai
Khoisan	(1)	Hottentot (Nama)

A.3 Inferring syllable structure from language descriptions. The generalizations of this study are based on a comparison of the syllable structures of the languages in the sample, as determined from a corpus of language descriptions. Complete normalization of the corpus is beyond my power, but, as in any comparative study, there is an obligation to make it as comparable as possible. The syllable structures cannot be taken directly from language descriptions, because the theoreticall viewpoints behind them are incompatible. The syllable structures of the corpus with which I work are tertiary data; they have been abstracted from language descriptions which are already abstractions of observations.

A.3.1 Conceptual framework for comparison. As a conceptual basis for interpreting language descriptions, I take the framework of generative phonology.[1] Syllable structures are compared at the systematic phonetic level. I assume that the segments in a string at this level are specified for the feature 'syllabic' and that the organization of the string into syllables is represented by specification of the adherence of each segment, i.e. what nucleus it 'belongs' to, roughly speaking. Since this level is considered to be the output of the phonological component of a grammar, phonological relations as well as phonetic properties are obviously relevant to its specification. Needless to say, none of the descriptions in the corpus characterize syllable structure in just this way. Nevertheless, in many cases the descriptions are easily interpretable in this framework. Others were not so straightforward. The usual problem I faced was that a number of studies assumed that either some particular aspect of phonetic form or else some underlying regularity determined syllable structure, and other possible relations were incompletely stated or not stated at all.

---

[1] A more detailed discussion of this framework and other theoretical bases for the description of syllable structure may be found in Bell (1970b, Chapter 2).

A.3.2 Guidelines for interpreting the data. The following examples illustrate the most commonly recurring or crucial cases where one must choose one of several possible ways to interpret the data available about a language's syllable structure.

Even though the basic level of reference is phonetic, phonemic descriptions did not usually pose a problem as long as sufficient phonetic detail was presented. It is still worth mentioning that considerable caution must be exercised against unusual phonemicizations. GUGU-YALANJI, an Australian language, (Oates 1964) provides an example. The syllable canon is claimed to be CV, CVC, CCV, CCVC. However, 'certain words of one phonetic syllable' containing a long vowel consist of two emic syllables, e.g. miyil [mi:l] 'eye', yawu [iau:] 'tortoise'. In the absence of further information one must add CVVC and CVV to the canon.

The criterion of minimum inventory, more than any other single factor, caused us to reinterpret descriptions. This was particularly troublesome for segmentation. Investigators sometimes justify their phonemic analysis by appealing to the minimization of the inventory of phonemes. Representations such as /hny/ for [ɲ] or [ɲɲ] are common, for example. Given either of these phonetic representations, I would count it as one segment. (A tautosyllabic sequence of marginal segments differing only in voicing is counted as a single segment.) Sufficient phonetic detail about clusters is often lacking, however.

This situation is met by counting complex segments as one segment, in the absence of explicit phonetic stipulation to the contrary. More specifically, two tautosyllabic marginal segments are counted together as one unit if one of the segments is interpretable as palatalization, velarization, pharyngealization, labialization, retroflexion, lateralization, affrication, aspiration, (de)voicing, or laryngealization. This means that, lacking contrary specification, [Nn], [Ll] and the like will

be one segment. (Note that no one would think of writing [Bb] for a 'voiceless unaspirated stop' in spite of its acoustic appropriateness.) The sequences /nt-/ and /-tn/ will be one segment but /mt-/ and /-tm/, two segments; /tl-/, one, but /pl-/, two; and so on.

A similar problem is posed for the characterization of the nucleus. Although I assume that the distinction represented by the feature '(nuclear) satellite' is not uncommon (Bailey 1968, 1969) and is potentially significant to the present study, it seems hopeless at this time to attempt to represent it systematically in our data. There is no consensus in the meaning of the 'imitation labels' available to describe this distinction, where it is noticed at all. The following arbitrary procedure will be followed (in the absence of clearly contrary phonetic description): postvocalic resonants will be counted as part of the final margin; postvocalic semivowels will be counted as part of the nucleus. In prevocalic position syllable-initial [y], [w], [i̯], [u̯], etc. count as marginal, unless otherwise stated. (Recall that tyV-, etc. may be treated as CV.)

The inventory of syllable types may also be minimized, a procedure explicitly advocated by Haugen (1965). This might have been a factor in the examples of GUGU-YALANJI, above, and ARABELA, below. The influence is sometimes explicitly acknowledged. Matteson and Pike (1958), for example, state that in PIRO, an Arawakan language of South America, 'the interpretation of /w/ and /j/ as consonants is based primarily on the rigid CV pattern of the language ... Clusters of unlike vowels do not occur.' The segment /w/ is said, however, to very freely to the position of [o] before consonants and front vowels, leading us to conclude that hiatus does occur, if only in free variation. Sommer's (1968, 1970) claim that the KUNJEN dialects of Australia have only closed syllables is an extreme case of economy overwhelming other considerations.

Descriptions which take the syllable to mean 'tone-bearing unit' are frequently incompatible with our framework. For example, Tucker (1962) calls the underlined segments of the following LUGANDA words 'syllabic' purportedly in a phonetic sense: /ǹmerè/, /èǹmerè/ 'food', /ɔ̀ppipa/ 'barrel'. Elsewhere he states that "the first part of the geminate [i.e. /pp/] is mere silence," and that "the long syllable [vowel + geminate stop or geminate nasal] ... in many respects behaves like two consecutive short syllables ... but ... in other respects its elements are too closely associated for it to be classed as a disyllable ..." The only evidence for the syllabicity of the segments in the words like the last two above is that they bear a tone in Tucker's system. My interpretation is that both the first two words are trisyllabic, the second having a closed first syllable.

A related difficulty arises when the syllable is conceived as an invariant of timing. In JAPANESE a unit of duration, the mora, has been frequently confused with the syllable.[2] Where a language's phonological processes compel treatment in terms of moras, I always distinguish mora from syllable.[3] The difficulty is not limited to mora-counting languages. In ARABELA, according to Rich (1963), "the syllable represents the basic unit of timing." Alerted for trouble, we are not disappointed. "The phonemes of /t n s r k/ have lengthened allophones which occur preceding homorganic consonant ... and

---

[2] Bloch (1950) was one linguist who used 'syllable' in this way for JAPANESE. It may have been a terminological convenience: "... the Japanese syllable is a unit of duration, such a unit is often called a mora; but the term 'syllable' is better established in descriptions of Japanese." (1950, 92)

[3] I take the mora to be unit of distance between two points in a string of segments. Moras are coherent with syllables; the length of a syllable is an integral number of moras -- either one or two. I further conjecture that the mora differs from the syllable in that it does not act as a domain; no phonological process requires as a context a mora boundary that is not simultaneously a syllable boundary.

constitute a phonemic syllable ..." Rich's sole criterion for syllabicity is that the segments' duration is roughtly that of other syllables. From the evidence presented, the straightforward interpretation, which I have adopted, is that these segments close the previous syllable. This is supported by the use of the term 'phonemic', suggesting an analysis contrary to the phone..c facts. Alternate forms indicate that the forms are the result of a recent vowel loss; further, they represent the only instances of a closed syllable in the language. Perhaps a concern to preserve an open syllable canon led to the given analysis.

Descriptions are of course found which adopt with a sonority-peak theory or a chest-pulse theory of the phonetic syllable. Neither can be maintained given our present knowledge.[4] Adherence to the first may give phonetic syllable structures incompatible with the level of phonetic representation adopted here. On the other hand, use of the label 'chest-pulse' is unlikely to indicate any incompatibility. There is no reason to suppose that in practice it was anything other than a completely neutral label corresponding to 'syllabic'.

Inevitably there remained unclear or incomplete descriptions for which there could be no general guide. The worst cases were those almost completely lacking phonetic detail. In the few instances found in the basic sample, the phonemic description was interpreted as if it were phonetic.

---

[4] Sonority-peak theories are discussed in Bell (1970, 31-39). The difficulties with chest-pulse theories are discussed by Ladefoged (1960) and in works cited therein.

BIBLIOGRAPHY

The following abbreviations are used:

AL	Anthropological Linguistics
ALS	African Language Studies
BSOAS	Bulletin of the School of Oriental and African Studies
IJAL	International Journal of American Linguistics
IL	Indian Linguistics
IURCAFL	Indiana University Research Center in Anthropology, Folklore, and Linguistics
JASA	Journal of the Acoustical Society of America
JAL	Journal of African Languages
JAOS	Journal of the American Oriental Society
JWAL	Journal of West African Languages
Lang.	Language
Ling. Inq.	Linguistic Inquiry
MIFAN	Memoir de l'Institut Francais de l'Afrique Noire
Misc. Ph.	Miscellanea Phonetica
NTS	Norsk Tideskrift for Sprogvidenskap
PICPS4	Proceedings of the Fourth International Congress of Phonetic Sciences, Helsinki, 4-9 September 1961. The Hague 1962.
POLA	Project on Linguistic Analysis. Reports. Second Series. Phonology Laboratory, University of California, Berkeley.
TPS	Transactions of the philological society
UAS	Uralic and Altaic Series
UCPL	University of California Publications in Linguistics
ZPSK	Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung

Abdel-Massih. 1968. Tamazight verb structure. A generative approach. Ann Arbor.

Abele, A. 1924. K voprosu o sloge. Slavia 3,1-34.

Abercrombie, David, et al., eds. 1964. In honour of Daniel Jones. London.

Abramson, Arthur S. 1962. The vowels and tones of standard Thai: acoustical measurements and experiments. IURCAFL 20. (= IJAL 28.2, Part II) Bloomington.

Allen, W.S. 1953. Phonetics in ancient India. London.

Allen, W.S. 1965. Vox Latina. A guide to the pronunciation of Classical Latin. Cambridge.

_____. 1968. Vox Graeca. A guide to the pronunciation of Classical Greek. Cambridge.

Andrzejewski, B.W. 1957. Some preliminary observations on the Borana dialect of Galla. BSOAS 19, 354-374.

Applegate, J.R. 1958. Outline of the structure of Shilha. Washington.

Aschmann, Herman P. 1946. Totonaco phonemes. IJAL 12, 34-43.

Avanesov, R.I. 1956. Fonetika sovremennogo russkogo literaturnogo jazyka. Moscow.

_____. 1968. Russkoe literaturnoe proiznoshenie. 4th ed. Moscow.

Bailey, C.-J. N. 1968. Syllabic; semi-syllabic; and unsyllabic in English: a ternary distinction PEGS Paper No. 22. February 20.

_____. 1969. Ternary-valued phonetic features: a third approximation. University of Hawaii Working papers in Linguistics 4, 63-80.

Bamgbose, Ayo. 1966. A grammar of Yoruba. West African Language Monographs 5. Cambridge.

Baptista, Priscilla and Ruth Wallin. 1968. Baure vowel elison. Linguistics 38, 5-11.

Barlow, A. Ruffell. 1960. Studies in Kikuyu grammar and idiom. Edinburgh.

Bauernschmidt, Amy. 1965. Amuzgo syllable dynamics. Lang. 41, 471-483.

Beach, D.M. 1938. The phonetics of the Hottentot language. Cambridge.

Bearth, Thomas and Hugo Zemp. 1967. The phonology of Dan (Santa) JAL 6, 9-29.

Bell, Alan. 1966. Machine syllabication of English speech. Final report project 188531-153. Stanford Research Institute, December 31.

Bell, Alan. 1970a. Syllabic consonants. Working Papers on Language Universals 4, B1-B49. Language Universals Project, Committee on Linguistics, Stanford University.

_____. 1970b. A state-process approach to syllabicity and syllable structure. Ph.D. Diss. Stanford University.

Bendor-Samuel, J.T. 1969. Yakur syllable patterns. Word 25, 16-23.

Biggs, Bruce. 1961. The structure of New Zealand Maaori. AL 3.3, 1-54.

Blanc, Haim. 1953. Studies in North Palestinian Arabic. Jerusalem.

Bloch, Bernard. 1950. Studies in colloquial Japanese IV, Phonemics. Lang. 26, 86-125.

Blood, David L. 1967. Phonological units in Cham. AL 9.8, 15-32.

Bogoras, Waldemar. 1922. Chuckchee. Handbook of American Indian languages, Part 2, Bulletin 40, Bureau of American Ethnology, ed. by Franz Boas, 633-903. Washington.

Borman, M.B. 1962. Cofan phonemes. In Elson (1962), 45-59.

Bôt Ba Njock, H.M. 1962. La déscription phonologique du Basa (Mbene), Thèse du doctorat. Université de Paris.

Bouquiaux. L. 1963. A propos de la phonologie du sara. JAL 3, 260-272.

Bowen, J. Donald. 1965. Beginning Tagalog. Berkeley.

Breatnach, Ristead B. 1947. The Irish of Ring Co. Waterford. A phonetic study. Dublin.

Burgess, Eunice and Patricia Ham. Multilevel conditioning of phoneme variants in Apinayé. Linguistics 41, 5-18.

Burling, Robbins. 1961. A Garo grammar. Poona.

_____. 1966. The addition of final stops in Maru. Lang. 42, 581-587.

_____. 1967. Proto Lolo-Burmese. IURCAFL 43 (= IJAL 33, 2. Part II) Bloomington.

Calame-Griaule, Geneviève. 1968. Dictionnaire dogon (dialecte tòr_) langue et civilisation. Paris.

Cardona, George. 1965. A Gujarati reference grammar. Philadelphia.

Carroll, Vern. No date. The structure of the language of Nukuoro. Wellington.

Chao, Yuen Ren. 1948. Mandarin primer. An intensive course in spoken Chinese. Cambridge.

_____. 1968. A grammar of spoken Chinese. Berkeley and Los Angeles.

Cheng, Teresa. 1968. The phonological system of Cantonese. POLA 5, C1-C85.

Coates, William A. and M.W.S. de Silva. 1960. The segmental phonemes of Sinhalese. University of Ceylon Review 18, 163-175.

Cohen, Marcel. 1936. Traité de langue amharique. Paris.

Cohen, Patrick D. 1966. Presyllables and reduplication in Jeh. In Thomas (1966), 31-40.

Cole, D.T. 1955. Introduction to Tswana grammar. London.

_____. 1967. Some features of Ganda linguistic structure. Johannesburg. (= African Studies 24, 3-54, 71-116, 119-240.)

Cooper, James and Nancy Cooper. 1966. Halàng phonemes. In Thomas (1966), 87-98.

Cornyn, W.S. 1945. Spoken Burmese. New York.

Cowan H.K.J. 1965a. Grammar of the Sentani language. Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 47. The Hague.

_____. 1965b. The Oirata language. Lingua 14, 360-370.

Crawford, John Chapman. 1963. Totontepec Mixe phonotagmemics. Norman.

Cubit, L.E. 1964. Kayan phonemics. Bijdragen tot de Taal-, Land- en Volkenkunde 120, 409-423.

Dirks, Sylvester. 1953. Campa (Arawak) phonemes. IJAL 19, 302-304.

Dixon, R.M.W. 1970. Olgolo syllable structure and what they are doing about it. Ling. Inq. 1, 273-276.

Dobson, E.J. 1968. English pronunciation 1500-1700. Vol. II Phonology. 2nd ed. Oxford.

Douglas, W.H. 1954-55. Phonology of the Australian aboriginal language spoken at Ooldea, South Australia, 1951-1952. Oceania 25, 216-229.

_____. 1964. An introduction to the Western Desert language. A pedagogical description of the Western Desert language, based on the dialect spoken at Warburton Ranges, Western Australia. (Revised.) Oceania Linguistic Monographs No. 4. Sydney.

_____. 1968. The aboriginal languages of the South-West of Australia. Australian Aboriginal Studies No. 14. Canberra.

Eckmann, János. 1962. The Turkish dialect of Edirne. In Poppe (1962), 45-69.

Elson, Benjamin, ed. 1962. Studies in Ecuadorian Indian languages: I. Norman.

_____. 1963. Studies in Peruvian Indian languages: I. Norman.

_____. and Comas. 1961. A William Cameron Townsend en el vigésimoquinto aniversario del Instituto Linguïstico de Verano. Mexico.

Emeneau, Murray B. 1955. Kolami. UCPL 12. Berkeley and Los Angeles.

Feller, William. 1968. An introduction to probability theory and its applications, Volume I, 3rd ed. New York.

Ferguson, Charles A. and Munier Chowdhury. 1960. The phonemes of Bengali. Lang. 36, 22-59.

Fouché, Pierre. 1959. Traité de prononciation francaise. Paris.

Fowler, Murray and Tasniya Isarasena. 1952. Total distribution of the sounds of Siamese: a topographical analysis. Madison.

Gedney, William J. 1965. Yay, a Northern Tai language in North Vietnam. Lingua 14, 180-193.

Gibson, Lorna F. 1956. Pame (Otomi) phonemics and morphophonemics. IJAL 22, 242-265.

Gradin, Dwight. 1966. Consonantal tone in Jeh phonemics. In Thomas (1966), 41-54.

Grammont, Maurice. 1933. Traité de phonetique. 8th ed. Paris.

Greenberg, Joseph H. 1941. Some problems in Hausa phonology. Lang. 17, 316-323.

_____. 1947. Swahili prosody. JAOS 67, 24-30.

_____. 1965. Some generalizations concerning initial and final consonant sequences. Linguistics 18, 5-34.

_____. 1966a. Synchronic and diachronic universals in phonology. Lang. 42, 508-517.

_____. 1966b. Universals of Language. 2nd ed. Cambridge.

_____. 1969. Some methods of dynamic comparison in linguistics. Substance and structure of language, ed. by Jaan Puhvel, 147-203. Berkeley and Los Angeles.

_____. To appear. The Indo-Pacific hypothesis. Current trends in linguistics. Vol. VIII, ed. by Thomas A. Sebeok. The Hague.

_____., Charles E. Osgood, and James J. Jenkins. 1966. Memorandum concerning language universals. In Greenberg (1966b), xv-xxvii.

Gregerson, Edgar. 1961. Luo. A grammar. Ph. D. Diss. Yale University.

Gregores, Emma and Jorge A. Suárez. 1967. A description of colloquial Guarani. The Hague.

Grimes, Joseph E. 1959. Huichol tone and intonation. IJAL 25, 221-232.

Hallaire, J. and J. Robinne. 1955-59. Dictionnaire sara-français. Mimeo. Koumra-Fourvière.

Hamp, Eric P. 1954. Componential restatement of syllable structure in Trique. IJAL 20, 206-209.

Harms, Robert T. 1962. Estonian grammar. UAS 12. Bloomington.

_____. 1966. Stress, voice, and length in Southern Paiute. IJAL 32, 228-235.

Harrell, Richard S. 1956. The phonology of colloquial Egyptian Arabic. Manuscript version of Ph. D. Diss. Harvard University.

_____., Laila Y. Tewfik and George D. Selim. 1963. Lessons in colloquial Egyptian Arabic. Revised ed. Washington.

Haslev, Marianne. 1967. The Maori stress system. Acta Linguistica Hafniensia 10, 211-216.

Haugen, Einar. 1956. The syllable in linguistic description. For Roman Jakobson, 213-221. The Hague.

Healey, Alan. 1964. Telefol phonology. Linguistic Circle of Canberra Publications Series B. No. 3. Canberra.

Hebert, Raymond J. 1962. Karačay phonology. In Poppe (1962), 97-113.

Henderson, E.J.A. 1952. The main features of Cambodian pronunciation. BSOAS 14, 149-174.

_____. 1965. Final k in Khasi: a secondary phonological pattern. Lingua 14, 459-466.

Henne, Henry. 1965. Sathewkok Hakka phonology. NTS 20, 109-161.

Henry, Jules. 1935. A Kaingang text. IJAL 8, 172-218.

Hiremath, R.C. 1961. The structure of Kannada. Dharwar.

Hockett, C.F. 1947. Peiping phonology. JAOS 67, 253-267.

Hodge, Carleton T. 1947. An outline of Hausa grammar. Language dissertation No. 41. Supplement to Lang. 23.4.

Hoff, B.J. 1968. The Carib language: phonology, morphonology, morphology, texts, and word index. Verhandlingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 55. The Hague.

Hoffman, Carl. 1963. A grammar of the Margi language. London.

Hohepa, Patrick W. 1967. A profile generative grammar of Maori. Indiana University Publications in Anthropology and Linguistics. Memoir 20 (= IJAL 33.2 Part II). Bloomington.

Holmer, Nils. 1966. Notes on the system of stress in Maori. Acta Linguistica Hafniensia 9, 163-172.

Holtved, Erik. 1952. Remarks on the Polar Eskimo dialect. IJAL 18, 20-24.

Houis, Maurice. 1963. Etude descriptive de la langue susu. MIFAN 67. Dakar.

Innes, Gordon. 1966. An introduction to Grebo. London.

Isačenko, A.V. 1947. Fonetika spisovnej Rustiny. Bratislava.

Jacob, J.M. 1960. The structure of the word in Old Khmer. BSOAS 23, 351-368.

_____. 1966. Some features of Khmer versification. In memory of J.R. Firth, ed. by C.E. Bazell et al., 227-247. London.

Jakobson, Roman and Morris Halle. 1956. Fundamentals of language. The Hague: Mouton.

Johnson, Orvill E. and Catherine Peeke. 1962. Phonemic units in the Secoya word. In Elson (1962), 78-95.

Johnstone, T.M. 1967a. Eastern Arabic dialect studies. London.

_____. 1967b. Aspects of syllabication in the spoken Arabic of 'Anaiza. BSOAS 30, 1-16.

Kachru, Braj. B. 1969. Kashmiri and other Dardic languages. In Sebeok (1969), 284-306.

Kakuk, Suzanne. 1958. Le dialect turc de Kazanlyk I, II. Acta Orientalia Hungarica 8, 167-187, 241-311.

Kalstrom, Marjorie R. and Eunice V. Pike. 1968. Stress in the phonological system of Eastern Popoloca. Phonetica 18, 16-30.

Kelkar, Ashok R. and Pran Nath Trisal. 1964. Kashmiri word phonology. AL 6.1, 13-22.

Kemeny, John G. and J. Laurie Shell. 1960. Finite Markov chains. Princeton: Van Nostrand.

Kensinger, Kenneth M. 1963. The phonological hierarchy of Cashinahua. In Elson (1963), 207-218.

Kervran, R.P.M. and R.P.M. Prost. 1969. Les parlers dogons. I. Donno so. Dakar.

Key, Harold and Mary Key. 1953. The phonemes of Sierra Nahuat. IJAL 19, 53-56.

Kisseberth, Charles W. 1970. On the functional unity of phonological rules. Ling. Inq. 1, 291-306.

Komárek, Miroslav. 1958. Historická mluvnice Ceská. Prague.

Kreinovič, E.A. 1958a. Yukagirskii jazyk. Moscow and Leningrad.

_____. 1958b. Opyt issledovanija struktury sloga v korjakskom Jazke. Kodlady i coobscenija instityta jazykoznanija AN SSSR. 11. 151-167.

Krupa, Viktor. 1966. Morpheme and word in Maori. Janua Linguarum. Series Practica 46. The Hague.

Kučera, H. 1961. The phonology of Czech. The Hague.

_____. and George K. Monroe. 1968. A comparative phonology of Russian, Czech, and German. New York.

Kuipers, Aert H. 1967. The Squamish language. Grammar, texts, dictionary. The Hague.

Kuryłowicz, Jerzy. 1948. Contribution à la théorie de la syllabe. Biuletyn Polskiego Towarzystwa Jezyko-Znawaczego 8, 80-114.

Ladefoged, Peter. 1958. Syllables and stress. Misc. Ph. 3, 1-14.

Lauritsen, Margaret. 1968. A phonetic study of the Danish stød. POLA 7. D1-D12.

Laver, John. 1967. A preliminary phonology of the Ayele dialect of Esako. JWAL 4.2, 53-56.

Lehn, Walter. 1963. Emphasis in Cairo Arabic. Lang. 39, 29-39.

Liccardi, Millicent and Joseph Grimes. 1968. Itonama intonation and phonemes. Linguistics 38, 36-41.

Lindskoog, John N. and Ruth M. Brend. 1962. Cayapa phonemics. In Elson (1962), 31-44.

Longacre, Robert E. 1952. Five phonemic pitch levels in Trique. Acta Linguistica Hafniensia 7, 62-82.

_____. 1955. Rejoinder to Hamp's Componential restatement of syllable structure in Trique. IJAL 21, 189-194.

_____. 1959. Trique tone morphemics. AL 4.4, 5-42.

Lopez, Cecilio. 1965. The Spanish overlay in Tagalog. Lingua 14, 467-504.

Luzbetak, Louis J. 1956. Middle Wahgi phonology. Oceania Linguistic Monographs No. 2. Sydney.

Lyovin, Anatole. 1968. Notes on the addition of final stops in Maru. POLA 7, L1-L22.

Mackenzie, D.N. 1958. Pashto verse. BSOAS 21, 319-333.

Martin, Samuel E. 1951. Korean phonemics. Lang. 27, 519-533.

_____. 1952. Morphophonemics of standard colloquial Japanese. Language dissertation No. 47. Baltimore (= Lang. 28.3 Part 2).

_____. 1959. Review of Japanische Phonetik von Günther Wenck. Lang. 35, 370-382.

Martinet, Andre. 1964. Economie des changements phonétiques. 2nd ed. Berne.

Maryott, K.R. 1961. The phonology and morphophonemics of Tabukang Sangir. Philippine Social Science and Humanities Review 26, 111-126.

Matisoff, James A. 1969. Review of Proto Lolo-Burmese by Robbins Burling. Lang. 44, 879-897.

Matteson, Esther. 1965. The Piro (Arawakan) language. UCPL 42. Berkeley and Los Angeles.

_____. and Kenneth L. Pike. 1958. Non-phonemic transition vocoids in Piro (Arawak). Misc. Ph. 3, 11-30.

Matthews, W. 1959. The phonematic system of literary Latvian. Rakstu Krajums. Veltijums Akademikim Profesoram Dr. Janim Endzelinam, ed. by E. Sokols et al., 181-220. Riga.

McArthur, Harry and Lucille McArthur. 1956. Aguacatec (Mayan) phonemes within the stress group. IJAL 22, 72-76.

McCawley, James E. 1965. The accentual system of standard Japanese. Ph. D. Thesis. M.I.T.

_____. 1968. The phonological component of a Japanese grammar. The Hague.

McIntosh, John B. 1945. Huichol phonemes. IJAL 11, 31-35.

McMahon, Ambrose. 1967. Phonemes and phonemic units of Cora (Mexico). IJAL 33, 128-134.

Meeussen, A.E. 1955. Les phonèmes du Ganda et du Bantou commun. Africa 25, 170-180.

Miller, Roy Andrew. 1935. Studies in spoken Tibetan, I: phonemics. JAOS 75, 40-51.

Miller, Wick R. 1965. Acoma grammar and texts. UCPL 40. Berkeley and Los Angeles.

Milner, G.B. 1958. Aspiration in two Polynesian languages. BSOAS 21, 368-375.

Mitchell, T.F. 1956. An introduction to Egyptian colloquial Arabic. London.

Mohr, Burckhard. 1970. Personal communication.

Moore, Bruce R. 1962. Correspondences in South Barbacoan Chibchan. In Elson (1962), 270-289.

Moreno, M.M. 1939. Grammatica teorico-practica della lingua Galla, con esercizi. Milan.

Morgenstierne, Georg. 1941. The phonology of Kashmiri. Acta Orientalia 19, 79-99.

Naes, Olav. 1965. Norsk grammatik. Elementaere strukturer og syntaks. 2nd ed. Oslo.

Navarro, Tomás T. 1957. Manual de pronunciación española. 6th ed. New York.

Nayak, H.M. 1967. Kannada. Literary and colloquial. A study of two styles. Mysore.

Newell, Leonard E. 1956. Phonology of the Gohang Ifugao dialect. Philippine Journal of Science 85, 523-539.

Noble, G. Kingsley. 1965. Proto-Arawakan and its descendants. IURCAFL 38 (= IJAL 31.3, part II). Bloomingtor and The Hague.

Noss, Richard B. 1964. Thai reference grammar. Washington.

_____. 1966. The treatment of */R/ in two modern Khmer dialects. Studies in comparative austroasiatic linguistics, ed. by Norman H. Zide, 88-95. (= Indo-Iranian Monographs 5). The Hague.

Oates, Lynette. 1964. Distribution of phonemes and syllables in Gugu-Yanlanji. AL 6.1, 23-26.

Oates, William and Lynette Oates. 1969. Gugu-Yalanji phonemes. Occasional Papers in Aboriginal Studies 2, 1-6.

O'Connor, J.D. and J.L.M. Trim. 1953. Vowel, consonant, syllable — a phonological definition. Word 9, 103-122.

O'Grady, G.N. 1957. Statistical investigations into an Australian language. Oceania 27, 283-312.

Olsen, Ronald D. 1967. The syllable in Chipaya. IJAL 33, 300-304.

Orr, Carolyn. 1962. Ecuador Quichua phonology. In Elson (1962), 60-77.

_____. 1962. Ecuador Quichua clause structure. In Elson (1962), 114-124.

Osborne, Henry A., Jr. 1966. Warao I: phonology and morphophonemics. IJAL 32, 108-123.

Ott, Willis and Rebecca Ott. 1967. Phonemes of the Ignaciano language. Linguistics 35, 56-70.

Palmer, F.R. 1957. The verb in Bilin. BSOAS 19, 131-159.

_____. 1958. The noun in Bilin. BSOAS 21, 376-391.

_____. 1960. An outline of Bilin phonology. Atti del convegno intern. di studi Etiopia, 109-116. Rome.

Pandit, P.B. 1957. Nasalization, aspiration and murmur in Gujarati. IL 17, 165-172.

_____. 1958. Duration, syllable and juncture in Gujarati. IL 19, 212-218. (Turner Jubilee Volume I)

Parzen, Emanuel. 1962. Stochastic processes. San Francisco.

Penzl, Herbert. 1955. A grammar of Pashto. Washington.

Petráček, K. 1965. Phonetik, Phonologie und Morphonologie der Berti- (Siga) Sprache in Dār Fur (Sūdān). Archiv Orientalni 33, 341-366.

Pickett, Velma. 1951. Nonphonemic stress: a problem in stress placement in Isthmus Zapotec. Word 7, 70-85.

_____. 1960. The grammatical hierarchy of Isthmus Zapotec Lang. 36.1 (part 2). 1-101. (= Language dissertation No. 56)

Pike, Eunice V. 1951. Tonemic-intonemic correlation in Mazahua (Otomí). IJAL 17, 37-41.

Pike, K.L. 1964. Stress trains in Auca. In Abercrombie (1964), 425-431.

_____. 1967. Language in relation to a unified theory of the structure of human behavior. 2nd ed. The Hague.

_____. and Eunice V. Pike. 1947. Immediate constituents of Mazateco syllables. IJAL 13, 78-91.

Pinnow, Heinz-Jurgen. 1959. Versuch einer historischen Lautlehre der Kharia-Sprache. Wiesbaden.

Pinnow, Heinz-Jurgen. 1964. Bemerkungen zur Phonetik und Phonemik des Kurukh. Indo-Iranian Journal 8, 32-59.

Pittman Richard S. 1961. The phonemes of Tetelcingo (Morelos) Nahuatl. In Elson and Comas (1961), 643-651.

Polomé, Edgar C. 1967. Swahili language handbook. Washington.

Poppe, Nicholas. 1964. Bashkir Manual. UAS 36. Bloomington.

_____. 1970. Mongolian language handbook. Washington.

Popperwell, R.G. 1963. The pronunciation of Norwegian. Cambridge and Oslo.

Prost, A. 1956. La langue soñay et ses dialectes. MIFAN 47. Dakar.

Prost, Gilbert R. 1967. Phonemes of the Chácobo language. Linguistics 35, 61-65.

Pulgram, Ernst. 1965. Consonant cluster, consonant sequence, and the syllable. Phonetica 13, 76-81.

_____. 1970 Syllable, word, nexus, cursus. The Hague: Mouton.

Rabel, Lili. 1961. Khasi, a language of Assam. Baton Rouge.

Rand, Earl. 1968. The structural phonology of Alabaman, a Muskogean language. IJAL 34, 94-103.

Redden, James E. 1966. Walapai I: phonology. IJAL 32, 1-16.

Rensch, Calvin. 1966. Proto-Otomangean phonology. University of Pennsylvania Ph. D. Diss.

_____. and Carolyn M. Rensch. 1966. The Lalana Chinantec syllable. Summa anthropologica en homenaje a Robert J. Weitlaner, ed. by Pompa y Pompa, 455-463. Mexico.

Reyburn, William D. 1954. Quechua I: phonemics. IJAL 20, 210-214.

Rich, Furne. 1963. Arabela phonemes and high-level phonology. In Elson (1963), 193-206.

Robbins, Frank. 1961. Quiotepec Chinantec syllable patterning. IJAL 27, 237-250.

Rowan, Orland. 1967. Phonology of Paresí (Arawakan). Acta Linguistica Hafniensia 10, 201-210.

Saint, Rachel and K. L. Pike. 1962. Auca phonemics. In Elson (1962), 2-30.

Salzmann, Zdenek. 1956. Arapahoe I: phonology. IJAL 22, 49-56.

Samarin, William J. 1966. The Gbeya language. Grammar, texts, and vocabularies. UCPL 44. Berkeley and Los Angeles.

Sanderson, G.M. 1954. A dictionary of the Yao language. Zomba.

Sangster, Linda and Emmanuel Faber. 1968. Susu basic course. Final report, Project No. 6-2320. U.S. Department of Health, Education, and Welfare. May.

Sapir, Edward. 1930. Southern Paiute, a Shoshonean language. Proceedings of the American Academy of Arts and Sciences 65, 1-3.

_____. and Harry Hoijer. 1967. The phonology and morphology of the Navaho language. UCPL 50. Berkeley and Los Angeles.

Sayers, Barbara. 1964. The phonemes of Coen Wik-Munkan. Occasional Papers in Aboriginal Studies 2, 51-56.

Schachter, Paul and Victoria Fromkin. 1968. A phonology of Akan: Akuapem, Asanti and Fante. Working Papers in Phonetics 9. UCLA Phonetics Laboratory.

Scott, N.C. 1948. A study in the phonetics of Fijian. BSOAS 12, 737-752.

Sebeok, Thomas A., ed. 1969. Current trends in linguistics. Volume 5, Linguistics in South Asia. New York.

Shapiro, Michael. 1968. Russian phonetic variants and phonostylistics. UCPL 49. Berkeley and Los Angeles.

Shetter, William Z. 1958. Phonemics of the Zwolle dialect — synchronic and diachronic. Lang. 34, 40-54.

Shevelov, George Y. and John J. Chew, Jr. 1969. Open syllable languages and their evolution: common Slavic and Japanese. Word 25, 252-274.

Shipley, William. 1956. The phonemes of Northeastern Maidu. IJAL 22, 233-237.

Shorto, H. L. 1960. Word and syllable patterns in Palaung. BSOAS 23, 552-553.

de Silva, M. W. S. 1959. Syllable structure in spoken Sinhalese: a prosodic statement. University of Ceylon Review 17, 106-116.

Sivertsen, Eva. 1968. Review of Vanvik, Arne: A phonetic-phonemic analysis of the dialect of Trondheim. NTS 22, 136-144.

Skinner, Leo E. 1962. Usila Chinantec syllable structure. IJAL 28, 251-255.

Skorik, P. Ja. 1961. Grammatika čukotskogo jazyka, čast' pervaja. Moscow-Leningrad.

Smalley, William S. 1961. Outline of Khmu? structure. New Haven.

Smith, Kenneth D. 1968. Laryngealization and de-laryngealization in Sedang phonemics. Linguistics 35, 52-69.

Sommer, Bruce A. 1969. Kunjen phonology: synchronic and diachronic. Pacific Linguistics Series B Monographs 11. Canberra.

_____. 1970. An Australian language without CV syllables. IJAL 36, 57-58.

Spotts, Hazel. 1953. Vowel harmony and consonant sequences in Mazahua (Otomi). IJAL 19, 253-258.

_____. 1956. Some post-conquest changes in Mazahua. IJAL 22, 208-211.

Stennes, Leslie H. 1967. A reference grammar of Adamawa Fulani. African Studies Center, Michigan State University. East Lansing.

Stevens, Alan M. 1968. Madurese phonology and morphology. American Oriental Series 52. New Haven.

Stewart, J. A. 1955. Manual of colloquial Burmese. London.

Swadesh, Morris. 1946. South Greenlandic (Eskimo). Linguistic structures of native America, ed. by Cornelius Osgood, 30-59. New York.

Swadesh, Morris. 1951. Kleinschmidt Centennial III: Unaaliq and Proto Eskimo. IJAL 17, 66-70.

_____. 1952a. Uraaliq and Proto Eskimo II: phonemics and morphophonemics. IJAL 18, 25-34.

_____. 1952b. Unaaliq and Proto Eskimo III. IJAL 18, 69-76.

_____. 1952c. Unaaliq and Proto Eskimo IV: diachronic notes. IJAL 18, 166-171.

Swift, Lloyd B. 1963. A reference grammar of Northern Turkish. UAS 19, Bloomington and The Hague.

Teeter, Karl V. 1957. The Wiyot language. UCPL 37. Berkeley and Los Angeles.

Thomas, David D., ed. 1966. Mon-Khmer studies III. Linguistic Circle of Saigon.

Tompa, József. 1968. Ungarische grammatik. The Hague.

Trnka, B. 1936. General laws of phonemic combinations. Travaux du cercle linguistique de Prague 6, 57-62.

Trubetzkoy, N.S. 1969. Principles of phonology. Berkeley and Los Angeles: University of California Press.

Tschenkéli, Kita. 1958. Einführung in die georgische Sprache. Vol. I. Zurich.

Tucker, A.N. 1940. The eastern Sudanic languages. Vol. I. London.

_____. 1946-7. Foreign sounds in Swahili. BSOAS 11, 854-871; 12, 214-232.

_____. 1962. The syllable in Luganda: a prosodic approach. JAL 1, 122-133.

_____. and U.J. Ashton. 1942. Swahili phonetics. African Studies 1, 77-103, 161-182.

Tung, Chao-hui. 1968. The phonological system of Gaoxiong, a Min dialect of Chinese. POLA 5, T1-T77.

Tyler, Stephen A. 1969. Koya: an outline grammar. UCPL 54. Berkeley and Los Angeles.

Uldall, H.J. 1954. Maidu phonetics. IJAL 20, 8-16.

Ulving, Tor. 1953. Consonant gradation in Eskimo. IJAL 19, 45-52.

Vansina, J. 1959. Esquisse de grammaire Bushong. Tervuren.

Vanvik, Arne. 1966. A phonetic-phonemic analysis of the dialect of Trondheim. Oslo.

Varma, Siddheshwar. 1964. Syllabication in the Kashmiri language. In Abercrombie (1964), 471-474.

Vorbichler, Anton. 1965. Die Phonologie und Morphonologie des Balese. Gluckstadt.

Ward, Ralph L. 1946. The loss of final consonants in Greek. Lang. 22, 102-108.

Watson, Richard. 1964. Pacoh phonemes. Mon-Khmer studies I, ed. by John E. Banker et al., 135-148. Saigon.

Welmers, William Everett. 1946. A descriptive grammar of Fanti, Language dissertation no. 39. (= Lang. 23.3 Suppl.) Baltimore.

Wenck, Gunther. 1954. Japanische Phonetik. Vol. 1. Wiesbaden.

Wheeler, Alva and Margaret Wheeler. 1962. Siona phonemics (Western Tucanoan). In Elson (1962), 96-111.

Whiteley, W.H. 1965. Notes on the Ci-mini dialect of Swahili. ALS 6, 67-72.

_____. 1966. A study of Yao sentences. Oxford.

Williamson, Kay. 1965. A grammar of the Kolokuma dialect of Ijo. West African Language Monograph Series 2. London.

Wiesemann, Ursula. 1964. Phonological syllables and words in Kaingang. Volkerkundliche Abhandlungen 1, 307-313.

Wolff, Hans. 1952. Osage I: phonemes and historical phonology. IJAL 18, 63-68.

Wonderley, William L. 1951. Zoque II; phonemes and morphology. IJAL 17, 105-123.

Woo, Nancy. 1970. Tone in Northern Tepehuan. IJAL 36, 18-30.

Zeps, Valdis J. 1962. Latvian and Finnic linguistic convergences. UAS 9. Bloomington.

Ziervogel, D. 1967. A handbook of the Zulu language. Pretoria.

_____. and R.S. Dan. 1961 A handbook of the Venda language. Pretoria.

Zlatoverxova, V.G. 1966. Fonetika birmanskogo literaturnogo jazyka. Moscow.